

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 12:38:09 ; Search time 80 Seconds
(without alignments)
816.462 Million cell updates/sec

Title: US-10-083-853B-1
Perfect score: 1704
Sequence: 1 MGDFMSKTPKAMATKAKID.....DIQLDHERCDICSSRPPR 317

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	678.5	39.8	641	4	Q14288	Q14288 homo sapien
2	678	39.8	1275	4	000362	000362 homo sapien
3	676.5	39.7	1275	4	000363	000363 homo sapien
4	674.5	39.6	1275	4	09Y5K0	09Y5K0 homo sapien
5	674.5	39.6	1275	4	08TE30	08TE30 homo sapien
6	674	39.6	712	4	Q14754	Q14754 homo sapien
7	673.5	39.5	1275	4	000370	000370 homo sapien
8	673.5	39.5	1275	4	000372	000372 homo sapien
9	672.5	39.5	1275	4	Q15604	Q15604 homo sapien
10	672.5	39.5	1275	4	Q12881	Q12881 homo sapien
11	672.5	39.5	1275	4	000360	000360 homo sapien
12	672.5	39.5	1275	4	000366	000366 homo sapien
13	672.5	39.5	1275	4	000368	000368 homo sapien
14	672.5	39.5	1275	4	Q9UN80	Q9UN80 homo sapien
15	672.5	39.5	1275	4	Q15606	Q15606 homo sapien
16	670.5	39.3	1275	4	000375	000375 homo sapien

17	661.5	38.8	1275	4	000378	000378 homo sapien
18	637	37.4	364	4	Q9BWM5	Q9BWM5 homo sapien
19	637	37.4	370	4	Q9GZP0	Q9GZP0 homo sapien
20	588	34.5	290	11	Q9D1L8	Q9D1L8 mus musculu
21	588	34.5	370	11	Q9Z517	Q9Z517 mus musculu
22	585	34.3	370	11	Q9EQ11	Q9EQ11 ractus norv
23	463.5	27.2	646	11	Q63779	Q63779 ractus norv
24	463.5	27.2	1300	11	P97692	P97692 ractus norv
25	425.5	25.0	275	11	Q63306	Q63306 ractus norv
26	425	24.9	1275	6	Q6Z558	Q6Z558 canis fami1
27	413.5	24.3	1252	11	Q9QWY2	Q9QWY2 mus musculu
28	413.5	24.3	1281	11	Q88913	Q88913 mus musculu
29	413.5	24.3	1281	11	Q88914	Q88914 mus musculu
30	413.5	24.3	1281	11	Q9QWY3	Q9QWY3 mus musculu
31	413.5	24.3	1281	11	Q9QWY0	Q9QWY0 mus musculu
32	413.5	24.3	1281	11	Q88915	Q88915 mus musculu
33	413.5	24.3	1281	11	Q9QW12	Q9QW12 mus musculu
34	413.5	24.3	1281	11	Q54850	Q54850 mus musculu
35	413.5	24.3	1300	11	Q08906	Q08906 mus musculu
36	412.5	24.2	1281	11	Q60713	Q60713 mus musculu
37	412.5	24.2	1281	11	Q91289	Q91289 mus musculu
38	412.5	24.2	1281	11	Q91288	Q91288 mus musculu
39	412.5	24.2	1300	11	Q61785	Q61785 mus musculu
40	409.5	24.0	1219	11	Q9QYM3	Q9QYM3 mus musculu
41	389.5	22.9	367	11	Q63778	Q63778 ractus norv
42	374	21.9	392	6	Q28360	Q28360 didelphis m
43	359	21.1	136	4	Q9HAG0	Q9HAG0 homo sapien
44	344	20.2	92	4	Q14269	Q14269 homo sapien
45	341	20.0	246	4	Q96G03	Q96G03 homo sapien

ALIGNMENTS

RESULT 1

ID Q14288 PRELIMINARY; PRT; 641 AA.
AC Q14288;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 75.3 kDa protein (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Woods-Samuels P.;
RL Submitted (APR-1989) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89233117; PubMed=2497061;
RA Woods-Samuels P.; Wong C.; Mathias S.L.; Scott A.F.;
RA Kazanian H.H. Jr.; Antonarakis S.E.;
RT "Characterization of a nondeleterious L1 insertion in an intron of the human factor VIII gene and further evidence of open reading frames in functional L1 elements";
RT Genomics 4:290-296(1989).
RL EMBL; M22334; AAA88038.1; -
DR InterPro; IPR000477; RVTse.
DR Pfam; PF00078; rvc; 1
KW Hypothetical protein; RNA-directed DNA polymerase.
FT NON_TER
FT 1
SQ SEQUENCE 641 AA; 75313 MW; 0A88757E7E3F66BF CRC64;

Query Match 39.8%; Score 678.5; DB 4; Length 641;

Best Local Similarity 71.8%; Pred. No. 6,9e+48;

Matches 148; Conservative 4; Mismatches 27; Indels 27; Gaps 5;
QY 1 MGDFMSKTPKAMATKAKIDKMDLILKLSFCTAKETITIVNQLTWEKIPATYSPDKGL 60
DB 356 VGRDFMSKTPKAMATKAKIDKMDLILKLSFCTAKETITIVNQLTWEKIPATYSPDKGL 415

```

QY 61 ISRIYNELKOIYKKKTKNPICKKWKDMNRHFSKGIYAAKKMKKSSLAIREMOIKTT 120
DB 416 ISRIYNELKOIYKKKTKNPICKKWKDMNRHFSKGIYAAKKMKKSSLAIREMOIKTT 475
QY 121 MRHYLTPVRMAIIKKSGNN---RDMDEAGNHHSQCTITRTKQTPHYLTHRW---1100S 174
DB 476 MRHYLTPVRMAIIKKSGNNRCMRGCEIG-----TLHCHWMDCKLVOP 518
QY 175 HWTV---LSDISELMHKTDRIVNLL 197
DB 519 LMKSVWFRLRDL-ELIEIFDPAIPL 543

RESULT 2
000362 PRELIMINARY; PRT; 1275 AA.
ID 000362;
AC 000362;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative p150.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97285120; PubMed=9140393;
RA Saasaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA Deberardinis R.J., Gabriel A., Swergold G.D., Kazanian H.H. Jr.;
RT "Many human L1 elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL; U93564; AAC51263.1; -
DR InterPro; IPR005135; Exo_endo_phos.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; rvt; 1.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1275 AA; 149201 MW; 23D51D6B4358F28 CRC64;

Query Match 39.8%; Score 678; DB 4; Length 1275;
Best Local Similarity 94.3%; Pred. No. 1.7e-47;
Matches 132; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 MGKDFMSKTPKAMATKATKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSDKGL 60
DB 990 VGKDFMSKTPKAMATKATKIDKMDLIKLSFCTAKETTRVNRQPTWEKIFATYSDKGL 1049
QY 61 ISRIYNELKOIYKKKTKNPICKKWKDMNRHFSKGIYAAKKMKKSSLAIREMOIKTT 120
DB 1050 ISRIYNELKOIYKKKTKNPICKKWKDMNRHFSKGIYAAKKMKKSSLAIREMOIKTT 1109
QY 121 MRHYLTPVRMAIIKKSGNNR 140
DB 1110 MRHYLTPVRMAIIKKSGNNR 1129

RESULT 3
000363 PRELIMINARY; PRT; 1275 AA.
ID 000363;
AC 000363;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative p150.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97285120; PubMed=9140393;

```

```

RA Saasaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA Deberardinis R.J., Gabriel A., Swergold G.D., Kazanian H.H. Jr.;
RT "Many human L1 elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL; U93565; AAC51264.1; -
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR000477; RVTse.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; rvt; 1.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1275 AA; 149008 MW; A866976A3FD8F74 CRC64;

Query Match 39.7%; Score 676.5; DB 4; Length 1275;
Best Local Similarity 71.8%; Pred. No. 2.3e-47;
Matches 148; Conservative 4; Mismatches 27; Indels 27; Gaps 5;

QY 1 MGKDFMSKTPKAMATKATKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSDKGL 60
DB 990 VGKDFMSKTPKAMATKATKIDKMDLIKLSFCTAKETTRVNRQPTWEKIFATYSDKGL 1049
QY 61 ISRIYNELKOIYKKKTKNPICKKWKDMNRHFSKGIYAAKKMKKSSLAIREMOIKTT 120
DB 1050 ISRIYNELKOIYKKKTKNPICKKWKDMNRHFSKGIYAAKKMKKSSLAIREMOIKTT 1109
QY 121 MRHYLTPVRMAIIKKSGNN---RDMDEAGNHHSQCTITRTKQTPHYLTHRW---1100S 174
DB 1110 MRHYLTPVRMAIIKKSGNNRCMRGCEIG-----TLHCHWMDCKLVOP 1152
QY 175 HWTV---LSDISELMHKTDRIVNLL 197
DB 1153 LMKSVWFRLRDL-ELIEIFDPAIPL 1177

RESULT 4
09Y5K0 PRELIMINARY; PRT; 1275 AA.
ID 09Y5K0;
AC 09Y5K0;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 149.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Divoky V., Indra K., Mrug M., Brabec V., Huisman T.H.J., Prchal J.T.;
RT "A novel mechanism of B-Chalasemia. The insertion of L1
RT retrotransposable element into B globin IVSII.";
RL Blood 88:148-148(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Kimberland M.L., Divoky V., Prchal J., Schwahn U., Berger W.,
RA Kazanian H.H.;
RT "Full-length human L1 insertions retain the capacity for high
RT frequency retrotransposition in cultured cells.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF149422; AAD38785.1; -
DR InterPro; IPR005135; Exo_endo_phos.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; rvt; 1.
KW Hypothetical protein, RNA-directed DNA polymerase.
SQ SEQUENCE 1275 AA; 149015 MW; 7E4AB526015ED37C CRC64;

Query Match 39.6%; Score 674.5; DB 4; Length 1275;
Best Local Similarity 71.4%; Pred. No. 3.3e-47;
Matches 147; Conservative 4; Mismatches 28; Indels 27; Gaps 5;

QY 1 MGKDFMSKTPKAMATKATKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSDKGL 60
DB 990 VGKDFMSKTPKAMATKATKIDKMDLIKLSFCTAKETTRVNRQPTWEKIFATYSDKGL 1049

```

```

QY 61 ISRIYNELKOIYKKTKNPDKKVVDMNRHFSKEGIYAAKKNKYSLSLAIREMOIKTT 120
DB 1050 ISRIYNELKOIYKKTKNNPIKKWAKDMNRHFSKEDIYAAKKNKYSLSLAIREMOIKTT 1109
QY 121 MRYHLTPVRMAIIKKSGNN---RDMDAGNHSOQITTRTKNOTPHVLTIRM---ILQOS 174
DB 1110 MRYHLTPVRMAIIKKSGNNRCWGCCEIG-----TLHGMWDCKLYOP 1152
QY 175 HWVTV---ISDISELMHKTDRIYNLL 197
DB 1153 LMKSVWRFLRDL-ELRIPPDPALPL 1177

RESULT 5
Q8TE30
ID 08TE30 PRELIMINARY; PRT; 1275 AA.
AC 08TE30;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Hypothetical 149.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21668188; PubMed=11810275;
RA Benjes S.M., Morris C.M.;
RT "A full-length and potentially active LINE element is integrated
RT polymorphically within the IGL locus in a genomically unstable region
RT of chromosome 32."
RL Hum. Genet. 109:628-637(2001).
DR EMBL; AF421375; AAL50637.1; -.
KW Hypothetical protein.
SQ SEQUENCE 1275 AA; 149010 MW; B327D9D50A581764 CRC64;

Query Match 39.6%; Score 674.5; DB 4; Length 1275;
Best Local Similarity 71.4%; Pred. No. 3.3e-47;
Matches 147; Conservative 4; Mismatches 28; Indels 27; Gaps 5;

QY 1 MGDFMSKTPKMATATKIDKMDLIKLSFCTAKETTRIVNROLTEWEXIFATYSPDKGL 60
DB 990 VGKDFMSKTPKMATATKIDKMDLIKLSFCTAKETTRIVNROLTEWEXIFATYSPDKGL 1049
QY 61 ISRIYNELKOIYKKTKNPDKKVVDMNRHFSKEGIYAAKKNKYSLSLAIREMOIKTT 120
DB 1050 ISRIYNELKOIYKKTKNNPIKKWAKDMNRHFSKEDIYAAKKNKYSLSLAIREMOIKTT 1109
QY 121 MRYHLTPVRMAIIKKSGNN---RDMDAGNHSOQITTRTKNOTPHVLTIRM---ILQOS 174
DB 1110 MRYHLTPVRMAIIKKSGNNRCWGCCEIG-----TLHGMWDCKLYOP 1152
QY 175 HWVTV---ISDISELMHKTDRIYNLL 197
DB 1153 LMKSVWRFLRDL-ELRIPPDPALPL 1177

RESULT 6
Q14754
ID 014754 PRELIMINARY; PRT; 712 AA.
AC 014754;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE ORF11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=PLACENTA;
RX MEDLINE=90332398; PubMed=2165587;
RA Hochjoh H., Minakami R., Sakaki Y.;
RT "Selective cloning of the Human LI (line-1) sequence which transposed
RT in a relatively recent past."
RL Nucleic Acids Res. 18:4099-4104(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=92319645; PubMed=1320255;
RA Minakami R., Kurase K., Etoh K., Furuhata Y., Hattori M., Sakaki Y.;
RT "Identification of an internal cis-element essential for the human LI
RT transcription and a nuclear factor(s) binding to the element."
RL Nucleic Acids Res. 20:3139-3145(1992).
DR EMBL; X52235; CAA36480.1; -.
DR InterPro; IPR000477; RVISE.
DR Pfam; PF00078; RVC1.1.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 712 AA; 83610 MW; 9123697871E4599F CRC64;

Query Match 39.6%; Score 674; DB 4; Length 712;
Best Local Similarity 93.6%; Pred. No. 1.8e-47;
Matches 131; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGDFMSKTPKMATATKIDKMDLIKLSFCTAKETTRIVNROLTEWEXIFATYSPDKGL 60
DB 427 MGDFMSKTPKMATATKIDKMDLIKLSFCTAKETTRIVNROLTEWEXIFATYSPDKGL 486
QY 61 ISRIYNELKOIYKKTKNPDKKVVDMNRHFSKEGIYAAKKNKYSLSLAIREMOIKTT 120
DB 487 ISRIYNELKOIYKKTKNNPIKKWAKDMNRHFSKEDIYAAKKNKYSLSLAIREMOIKTT 546
QY 121 MRYHLTPVRMAIIKKSGNNR 140
DB 547 MRYHLTPVRMAIIKKSGNNR 566

RESULT 7
Q00370
ID 000370 PRELIMINARY; PRT; 1275 AA.
AC 000370;
DT 01-JUL-1997 (TEMBLrel. 04, Created)
DT 01-JUL-1997 (TEMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Putative p150.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97285120; PubMed=9140393;
RA Sassaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human LI elements are capable of retrotransposition."
RL Nat. Genet. 16:37-43(1997).
DR EMBL; U93569; AAC51271.1; -.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR000477; RVISE.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; RVC1.1.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1275 AA; 149011 MW; 588703688E7129FF CRC64;

Query Match 39.5%; Score 673.5; DB 4; Length 1275;
Best Local Similarity 71.4%; Pred. No. 4e-47;
Matches 147; Conservative 4; Mismatches 28; Indels 27; Gaps 5;

QY 1 MGDFMSKTPKMATATKIDKMDLIKLSFCTAKETTRIVNROLTEWEXIFATYSPDKGL 60
DB 990 VGKDFMSKTPKMATATKIDKMDLIKLSFCTAKETTRIVNROLTEWEXIFATYSPDKGL 1049
QY 61 ISRIYNELKOIYKKTKNPDKKVVDMNRHFSKEGIYAAKKNKYSLSLAIREMOIKTT 120

```

```

Db 1050 ISRIYNELKQIYKKKTNNPIKKAQKDMNRHFSKEDIYAAKGMKKCSSLAIREMOIKTT 1109
Qy 121 MRHYLTPVRMAIIKKSGNN--RDMDEAGNHSSQOTITRTKQTPHYLTHRW---110QS 174
Db 1110 MRHYLTPVRMAIIKKSGNNRCWRGCEIG-----TLHGMWCKLVOP 1152
Qy 175 HWTV---LSDISELMHKTDRIVNLL 197
Db 1153 LKSVWRFLDL-ELFIPDPAILL 1177

```

RESULT 8

```

000372 ID 000372 PRELIMINARY; PRT; 1275 AA.
AC 000372;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative p150.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97285120; PubMed=9140393;
RA Sasaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA Debernardi R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human L1 elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL: U93570; AAC51273.1;
DR InterPro: IPR005135; Exo_endo_phos.
DR InterPro: IPR004477; RVTse.
DR Pfam: PF03372; Exo_endo_phos. 1.
DR Rfam: RF00078; rvt. 1.
KM RNA-directed DNA polymerase.
SQ SEQUENCE 1275 AA; 149111 MW; 4711B3BC22F674E CRC64;

```

```

Query Match 39.5%; Score 673.5; DB 4; Length 1275;
Best Local Similarity 71.4%; Pred. No. 4e-47;
Matches 146; Conservative 5; Mismatches 31; Indels 21; Gaps 4;
Qy 1 MGKDFMSKTPKAMATKAKIDKMDILKLSFCTAKETITRVNROLTEWEXIFATYSPDKL 60
Db 990 VGKDFMSKTPKAMATKAKIDKMDILKLSFCTAKETITRVNROLTEWEXIFATYSPDKL 1049
Qy 61 ISRIYNELKQIYKKKTNNPIKKAQKDMNRHFSKEDIYAAKGMKKCSSLAIREMOIKTT 120
Db 1050 ISRIYNELKQIYKKKTNNPIKKAQKDMNRHFSKEDIYAAKGMKKCSSLAIREMOIKTT 1109
Qy 121 MRHYLTPVRMAIIKKSGNNRCWRGCEIG-----TLHGMWCKLVOP 1155
Db 1110 MRHYLTPVRMAIIKKSGNNRCWRGCEIG-----TLHGMWCKLVOP 1155
Qy 178 TV---LSDISELMHKTDRIVNLL 197
Db 1156 SVWRFLDL-ELFIPDPAILL 1177

```

RESULT 9

```

015604 ID 015604 PRELIMINARY; PRT; 1275 AA.
AC 015604;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE ORF2 contains a reverse transcriptase domain.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Dombroski B.A.;
RT "Isolation of an active human transposable element.";
RL Science 0:0-0(1991).
DR EMBL: M80340; AA51622.1;
DR InterPro: IPR005135; Exo_endo_phos.
DR InterPro: IPR004477; RVTse.
DR Pfam: PF03372; Exo_endo_phos. 1.
DR Rfam: RF00078; rvt. 1.
KM RNA-directed DNA polymerase.
SQ SEQUENCE 1275 AA; 149044 MW; C090473BBB0A44C1 CRC64;

```

Query Match

```

39.5%; Score 672.5; DB 4; Length 1275;
Best Local Similarity 71.4%; Pred. No. 4.8e-47;
Matches 147; Conservative 4; Mismatches 28; Indels 27; Gaps 5;
Qy 1 MGKDFMSKTPKAMATKAKIDKMDILKLSFCTAKETITRVNROLTEWEXIFATYSPDKL 60
Db 990 VGKDFMSKTPKAMATKAKIDKMDILKLSFCTAKETITRVNROLTEWEXIFATYSPDKL 1049
Qy 61 ISRIYNELKQIYKKKTNNPIKKAQKDMNRHFSKEDIYAAKGMKKCSSLAIREMOIKTT 120
Db 1050 ISRIYNELKQIYKKKTNNPIKKAQKDMNRHFSKEDIYAAKGMKKCSSLAIREMOIKTT 1109
Qy 121 MRHYLTPVRMAIIKKSGNN--RDMDEAGNHSSQOTITRTKQTPHYLTHRW---110QS 174
Db 1110 MRHYLTPVRMAIIKKSGNNRCWRGCEIG-----TLHGMWCKLVOP 1152
Qy 175 HWTV---LSDISELMHKTDRIVNLL 197
Db 1153 LKSVWRFLDL-ELFIPDPAILL 1177

```

RESULT 10

```

012881 ID 012881 PRELIMINARY; PRT; 1275 AA.
AC 012881;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Retrotransposible L1 element LRE2 from chromosome 1q.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TRANSPOSON=LRE2;
RA Holmes S.E., Dombroski B.A., Krebs C.M., Boehm C.D., Kazazian H.H.;
RT "A new retrotransposable human L1 element from the LRE2 locus on
RT chromosome 1q produces a chimaeric insertion.";
RL Nat. Genet. 7:143-148(1994).
DR EMBL: U09116; AAB60345.1;
DR InterPro: IPR005135; Exo_endo_phos.
DR InterPro: IPR004477; RVTse.
DR Pfam: PF03372; Exo_endo_phos. 1.
DR Rfam: RF00078; rvt. 1.
KM RNA-directed DNA polymerase.
FT VARIANT 1182 1182 E -> K.
SQ SEQUENCE 1275 AA; 149064 MW; CE4429261FEE223 CRC64;

```

Query Match

```

39.5%; Score 672.5; DB 4; Length 1275;
Best Local Similarity 71.4%; Pred. No. 4.8e-47;
Matches 147; Conservative 4; Mismatches 28; Indels 27; Gaps 5;
Qy 1 MGKDFMSKTPKAMATKAKIDKMDILKLSFCTAKETITRVNROLTEWEXIFATYSPDKL 60
Db 990 VGKDFMSKTPKAMATKAKIDKMDILKLSFCTAKETITRVNROLTEWEXIFATYSPDKL 1049
Qy 61 ISRIYNELKQIYKKKTNNPIKKAQKDMNRHFSKEDIYAAKGMKKCSSLAIREMOIKTT 120
Db 1050 ISRIYNELKQIYKKKTNNPIKKAQKDMNRHFSKEDIYAAKGMKKCSSLAIREMOIKTT 1109

```

```
QY 121 MRYHLTPVMAIIKKSNN---RDMDAGNHSQOTITRTKNQTPHYLTHRW---ILQOS 174
DB 1110 MRYHLTPVMAIIKKSNNRCRCGCEIG-----TLHCWMDCKLVOP 1152
QY 175 HWTV---LSDISELMHKTDRIVNLL 197
DB 1153 LKSWVRFLRDL-ELFIPDPALPL 1177

RESULT 11
000360 PRELIMINARY; PRT; 1275 AA.
ID 000360
AC 000360;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative p150.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=97285120; PubMed=9140393;
RA Sassaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human LI elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL; U93563; AAC51261.1;
DR InterPro; IPR005135; Exo_endo_phos.
DR pfam; PF03372; Exo_endo_phos; 1.
DR pfam; PF00078; rvc; 1.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1275 AA; 149062 MW; 350B4F0358E525F0 CRC64;

Query Match 39.5%; Score 672.5; DB 4; Length 1275;
Best Local Similarity 71.4%; Pred. No. 4.8e-47;
Matches 147; Conservative 3; Mismatches 28; Indels 27; Gaps 5;

QY 2 GQDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 61
DB 991 GQDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 1050
QY 62 SRIYNELKOIYKKTKNPIKAVYDMNRHPSKEGIYAAKKMKYSSSLAIRMOIKTTM 121
DB 1051 SRIYNELKOIYKKTKNPIKAVYDMNRHPSKEGIYAAKKMKYSSSLAIRMOIKTTM 1110
QY 122 RYHLTEVMAIIKKSNN---RDMDAGNHSQOTITRTKNQTPHYLTHRW---ILQOSH 175
DB 1111 RYHLTEVMAIIKKSNNRCRCGCEIG-----TLHCWMDCKLVOP 1153
QY 176 WTVV---LSDISELMHKTDRIVNLL 197
DB 1154 WKSVMRFLRDL-ELFIPDPALPL 1177

RESULT 12
000366 PRELIMINARY; PRT; 1275 AA.
ID 000366
AC 000366;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative p150.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=97285120; PubMed=9140393;
```

```
RA Sassaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human LI elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL; U93567; AAC51267.1;
DR InterPro; IPR005135; Exo_endo_phos.
DR pfam; PF03372; Exo_endo_phos; 1.
DR pfam; PF00078; rvc; 1.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1275 AA; 149078 MW; 3BB3BD2C2E06B61 CRC64;

Query Match 39.5%; Score 672.5; DB 4; Length 1275;
Best Local Similarity 71.4%; Pred. No. 4.8e-47;
Matches 147; Conservative 4; Mismatches 28; Indels 27; Gaps 5;

QY 1 MGDFFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 60
DB 990 MGDFFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 1049
QY 61 ISRIYNELKOIYKKTKNPIKAVYDMNRHPSKEGIYAAKKMKYSSSLAIRMOIKTT 120
DB 1050 ISRIYNELKOIYKKTKNPIKAVYDMNRHPSKEGIYAAKKMKYSSSLAIRMOIKTT 1109
QY 121 MRYHLTPVMAIIKKSNN---RDMDAGNHSQOTITRTKNQTPHYLTHRW---ILQOS 174
DB 1110 MRYHLTPVMAIIKKSNNRCRCGCEIG-----TLHCWMDCKLVOP 1152
QY 175 HWTV---LSDISELMHKTDRIVNLL 197
DB 1153 LKSWVRFLRDL-ELFIPDPALPL 1177

RESULT 13
000368 PRELIMINARY; PRT; 1275 AA.
ID 000368
AC 000368;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative p150.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=97285120; PubMed=9140393;
RA Sassaman D.M., Dombroski B.A., Moran J.V., Kimberland M.L., Naas T.P.,
RA DeBerardinis R.J., Gabriel A., Swergold G.D., Kazazian H.H. Jr.;
RT "Many human LI elements are capable of retrotransposition.";
RL Nat. Genet. 16:37-43(1997).
DR EMBL; U93568; AAC51269.1;
DR InterPro; IPR005135; Exo_endo_phos.
DR pfam; PF03372; Exo_endo_phos; 1.
DR pfam; PF00078; rvc; 1.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1275 AA; 149078 MW; 3BB3BD2C2E06B61 CRC64;

Query Match 39.5%; Score 672.5; DB 4; Length 1275;
Best Local Similarity 71.4%; Pred. No. 4.8e-47;
Matches 147; Conservative 4; Mismatches 28; Indels 27; Gaps 5;

QY 1 MGDFFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 60
DB 990 MGDFFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 1049
QY 61 ISRIYNELKOIYKKTKNPIKAVYDMNRHPSKEGIYAAKKMKYSSSLAIRMOIKTT 120
DB 1050 ISRIYNELKOIYKKTKNPIKAVYDMNRHPSKEGIYAAKKMKYSSSLAIRMOIKTT 1109
QY 121 MRYHLTPVMAIIKKSNN---RDMDAGNHSQOTITRTKNQTPHYLTHRW---ILQOS 174
```

```

Db 1110 MRVHLPVRAAIKKSGNNRCWCGEIG-----TLHGMWCKLVOP 1152
Qy 175 HWVTV---LSDISELMHKTDRIVNLL 197
Db 1153 LMKSVWRFLRDL-ELEIPDPALPIL 1177

```

RESULT 14

```

Q9UN80 PRELIMINARY; PRT; 1275 AA.
ID Q9UN80
AC Q9UN80;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical 149.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=RETROTRANSPOSON L1;
RX MEDLINE=98361157; PubMed=9697692;
RA Schwann U., Lenzner S., Dong J., Feil S., Hinzmann B.,
RA van Dijkhoven G., Kirschner R., Hemberger M., Bergen A.A.,
RA Rosenberg T., Pluckers A.J., Fundele R., Rosenthal A., Cremers F.P.,
RA Roper H.H., Berger W.,
RT "Positional cloning of the gene for X-linked retinitis pigmentosa 2.",
RL Nat. Genet. 19:327-332(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=RETROTRANSPOSON L1;
RA Kimberland M.L., Divoky V., Prechal J., Schwann U., Berger W.,
RA Kazazian H.H.;
RT "Full-length human L1 insertions retain the capacity for high
RT frequency retrotransposition in cultured cells.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AFI48856; AAD39215.1; -
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR004777; RVTse.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; RVC; 1.
KM Hypothetical protein; RNA-directed DNA polymerase.
SQ SEQUENCE 1275 AA; 149043 MW; A93622F87A62CE4 CRC64;

```

Query Match 39.5%; Score 672.5; DB 4; Length 1275;

Best Local Similarity 71.4%; Pred. No. 4.8e-47;

Matches 147; Conservative 4; Mismatches 28; Indels 27; Gaps 5;

```

Qy 1 MGKDFMSKTPKAMATKAKIDKMDLILKSFCTAKETIRVNRQLTWEKIFATYSPDKGL 60
Db 990 VGKDFMSKTPKAMATKAKIDKMDLILKSFCTAKETIRVNRQLTWEKIFATYSPDKGL 1049
Qy 61 ISRIYNEIKQIYKKKTNNPIKKWVKDNNRHFSEKGIYAAKGMKYSLSAIREMOIKTT 120
Db 1050 ISRIYNEIKQIYKKKTNNPIKKWVKDNNRHFSEKEDIYAAKGMKCCSSLSAIREMOIKTT 1109
Qy 121 MRVHLPVRAAIKKSGNN---RMDDEAGNHSSQOTTITRTKNOTPHVLTIRW---ILQOS 174
Db 1110 MRVHLPVRAAIKKSGNNRCWCGEIG-----TLHGMWCKLVOP 1152
Qy 175 HWVTV---LSDISELMHKTDRIVNLL 197
Db 1153 LMKSVWRFLRDL-ELEIPDPALPIL 1177

```

RESULT 15

```

ID Q15606 PRELIMINARY; PRT; 1275 AA.
AC Q15606;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)

```

```

DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Hypothetical 149.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dombroski B.A.;
RL Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 84-1275 FROM N.A.
RA Woods-Samuels P.;
RL Submitted (APR-1989) to the EMBL/GenBank/DBJ databases.
DR EMBL; M80343; AAB59368.1; -
DR EMBL; M22333; AAB8037.1; -
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR004777; RVTse.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Pfam; PF00078; RVC; 1.
KM Hypothetical protein; RNA-directed DNA polymerase.
SQ SEQUENCE 1275 AA; 148971 MW; 6DA1229742848323 CRC64;

```

Query Match 39.5%; Score 672.5; DB 4; Length 1275;

Best Local Similarity 71.4%; Pred. No. 4.8e-47;

Matches 147; Conservative 4; Mismatches 28; Indels 27; Gaps 5;

```

Qy 1 MGKDFMSKTPKAMATKAKIDKMDLILKSFCTAKETIRVNRQLTWEKIFATYSPDKGL 60
Db 990 VGKDFMSKTPKAMATKAKIDKMDLILKSFCTAKETIRVNRQLTWEKIFATYSPDKGL 1049
Qy 61 ISRIYNEIKQIYKKKTNNPIKKWVKDNNRHFSEKGIYAAKGMKYSLSAIREMOIKTT 120
Db 1050 ISRIYNEIKQIYKKKTNNPIKKWVKDNNRHFSEKEDIYAAKGMKCCSSLSAIREMOIKTT 1109
Qy 121 MRVHLPVRAAIKKSGNN---RMDDEAGNHSSQOTTITRTKNOTPHVLTIRW---ILQOS 174
Db 1110 MRVHLPVRAAIKKSGNNRCWCGEIG-----TLHGMWCKLVOP 1152
Qy 175 HWVTV---LSDISELMHKTDRIVNLL 197
Db 1153 LMKSVWRFLRDL-ELEIPDPALPIL 1177

```

Search completed: June 5, 2003, 12:43:48

Job time : 82 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 12:39:49 / Search time 40 Seconds
(without alignments)
761.865 Million cell updates/sec

Title: US-10-083-853B-1
Perfect score: 1704
Sequence: 1 MGDFMSKTPKAMATKAKID.....DQLDHERCDICSSRPPR 317

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: PIR 73:1*
2: PIR1:1*
3: PIR2:1*
4: PIR3:1*
5: PIR4:1*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	679.5	39.9	1280	2 B34087	hypothetical prote
2	674	39.6	712	2 S23650	retrovirus-related
3	672.5	39.5	1275	2 I38588	reverse transcript
4	672.5	39.5	1275	2 S65824	reverse transcript
5	668.5	39.2	1275	2 B28096	line-1 protein ORF
6	658.5	38.6	1259	4 GNHUI1	retrovirus-related
7	637	37.4	370	2 JC7581	spinal cord-derive
8	585	34.3	370	2 JC7582	spinal cord-derive
9	443.5	26.0	1260	4 GNLRL1	retrovirus-related
10	425.5	25.0	275	2 S21348	probable pol polyp
11	412.5	24.2	1281	1 GNSML	retrovirus-related
12	374	21.9	392	2 S57662	hypothetical prote
13	344	20.2	92	2 F41935	hypothetical prote
14	343.5	20.2	562	2 JU0033	hypothetical Li pr
15	271	15.9	513	2 S21976	probable RNA-dirac
16	243.5	14.3	143	2 E41925	hypothetical prote
17	125.5	7.4	148	2 D49530	16k vascular endoc
18	101	5.9	1895	2 T15881	hypothetical prote
19	100.5	5.9	625	2 H70330	hypothetical prote
20	98	5.8	4550	2 T18440	hypothetical prote
21	96.5	5.7	419	2 S69207	vascular endotheli
22	95.5	5.6	188	2 JC4679	vascular endotheli
23	95.5	5.6	207	2 JC4679	vascular endotheli
24	95	5.6	2562	2 T14266	kin protein, - chic
25	94	5.5	166	2 JN0248	vascular-derived g
26	94	5.5	188	2 S80735	platelet-derived g
27	92.5	5.4	870	2 GS6450	PSD1.31 protein -
28	92.5	5.4	959	2 C82392	oxidoreductase/iro
29	92.5	5.4	4092	1 S38128	dynein heavy chain

30	92	5.4	133	2 B49530	vascular endotheli
31	91.5	5.4	270	2 T47421	hypothetical prote
32	91.5	5.4	4668	2 F82885	hypothetical prote
33	91	5.3	803	2 B84931	DNA topoisomerase
34	90.5	5.3	344	2 T20209	hypothetical prote
35	90.5	5.3	357	2 B98056	hypothetical prote
36	90.5	5.3	739	2 E86434	protein F178.27 f
37	90.5	5.3	1817	2 D71606	hypothetical prote
38	90.5	5.3	2437	2 T18482	hypothetical prote
39	90	5.3	849	2 H84668	Mutator-like trans
40	89.5	5.3	357	2 A95190	membrane associate
41	89.5	5.3	455	2 H97237	DBP4 protein - Yea
42	89.5	5.3	704	2 S25371	hypothetical prote
43	89.5	5.3	2539	2 B71619	hypothetical prote
44	89	5.2	196	2 A37359	platelet-derived g
45	89	5.2	196	2 B28964	platelet-derived g

ALIGNMENTS

RESULT 1
B34087
hypothetical protein (L1H 3' region) - human
C/Species: Homo sapiens (man)
C/Date: 30-Mar-1990 #sequence_revision 30-Mar-1990 #text_change 30-Sep-1993
C/Accession: B34087
R/Scott, A.F.; Schmeckpeper, B.J.; Abdelrazik, M.; Comey, C.T.; O'Hara, B.; Rossiter, J.I
Genomics 1, 113-125, 1987
A/Title: Origin of the human L1 elements: proposed progenitor genes deduced from a conser
A/Reference number: A34087; PMID:88085155; PMID:3692483
A/Accession: B34087
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-1280 <SCO>
C/Superfamily: pol polyprotein

Query Match 39.9%; Score 679.5; DB 2; Length 1280;
Best Local Similarity 71.8%; Pred. No. 4.5e-45;
Matches 148; Conservative 3; Mismatches 28; Indels 27; Gaps 5;

QY 1 MGDFMSKTPKAMATKAKIDKMDLILKKSFCYKETTIRVNRQLTWEKIFATYSPDKL 60
DB 995 MGDFMSKTPKAMATKAKIDKMDLILKKSFCYKETTIRVNRQLTWEKIFATYSPDKL 1054
QY 61 ISSIYNELKQIYKKTNPDKVWVKNRHRFSKGIYAAKQNKYSSSLAIREMOIKTT 120
DB 1055 ISSIYNELKQIYKKTNPDKVWVKNRHRFSKGIYAAKQNKYSSSLAIREMOIKTT 1114
QY 121 MRVHLTPVNRALIKKSGNN---RDMDEAGNHHSQQTITRTKQTPHVLTHRW---ILQOS 174
DB 1115 MRVHLTPVNRALIKKSGNNRNCWGCCEIG-----TLHLCWDCCKLQVP 1157
QY 175 HWTVV---LSDISLNAKHTDRIYNLL 197
DB 1158 LWKTIVRFLKDL-ELSLPDPALPL 1182

RESULT 2
S23650
retrovirus-related hypothetical protein II - human retrotransposon LINE-1
C/Species: Homo sapiens (man)
C/Date: 22-Nov-1993 #sequence_revision 01-Nov-1996 #text_change 08-Jan-1999
C/Accession: S23650
R/Hohjoh, H.; Minakami, R.; Sakaki, Y.
Nucleic Acids Res. 18, 4093-4104, 1990
A/Title: Selective cloning and sequence analysis of the human L1 (LINE-1) sequences which
A/Reference number: S23649; PMID:9032338; PMID:2165587
A/Accession: S23650
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-712 <HOH>
A/Cross-references: EMBL:X52235

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1990

C/Genetics:

A/Mobile element: LINE-1

A/Start codon: GTG

C/Superfamily: pol polyprotein

Query March 39.6%; Score 674; DB 2; Length 712;
Best Local Similarity 93.6%; Pred. No. 6.2e-45;
Matches 131; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 MGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 60
DB 427 MGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 486
QY 61 ISRIYNELKOIYKKKTNPPIKKWVDMNRHFSKEGIYAAKGMKKYSSSLAIREMOIKTT 120
DB 487 ISRIYNELKOIYKKKTNPPIKKWVDMNRHFSKEDIYAAKGMKKYSSSLAIREMOIKTT 546
QY 121 MRVHLTPVRMAIIKKSNNR 140
DB 547 MRVHLTPVRMAIIKKSNNR 566

RESULT 3
138588
reverse transcriptase homolog - human retrotransposon L1

N/Alternate names: ORF2 protein

C/Species: Homo sapiens (man)

C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 26-Aug-1999

C/Accession: I38588

R/Holmes, S.E.; Dombroski, B.A.; Krebs, C.M.; Boehm, C.D.; Kazanian, H.H.

Nature Genet. 7, 143-148, 1994

A/Title: A new retrotransposable human L1 element from the LBR2 locus on chromosome 1q F

A/Reference number: I38587; MUID:95004577; PMID:7920631

A/Accession: I38588

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1275 <RES>

A/Cross-references: EMBL:U09116; NID:9483914; PIDN:AA60345.1; PID:9483916

C/Superfamily: pol polyprotein

Query Match 39.5%; Score 672.5; DB 2; Length 1275;
Best Local Similarity 71.4%; Pred. No. 1.6e-44;
Matches 147; Conservative 4; Mismatches 28; Indels 27; Gaps 5;

QY 1 MGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 60
DB 990 VGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 1049
QY 61 ISRIYNELKOIYKKKTNPPIKKWVDMNRHFSKEGIYAAKGMKKYSSSLAIREMOIKTT 120
DB 1050 ISRIYNELKOIYKKKTNPPIKKWVDMNRHFSKEDIYAAKGMKKYSSSLAIREMOIKTT 1109
QY 121 MRVHLTPVRMAIIKKSNNR--RDMDKGNHHSQOTTRTKNOTPHVLTTRW---ILQOS 174
DB 1110 MRVHLTPVRMAIIKKSNNRCMRGCGEIG-----TLHGMWCKLVOP 1152
QY 175 HWYTV---ISDISELMHKTDRIVNLL 197
DB 1153 LMKSVWRFRLDL-ELRIPDPALPLL 1177

RESULT 4

S65824
reverse transcriptase homolog - human transposon L1.1

C/Species: Homo sapiens (man)

C/Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 26-Aug-1999

C/Accession: S65824

R/Dombroski, B.A.

submitted to the EMBL Data Library, January 1992

A/Description: Isolation of an active human transposable element.

A/Reference number: S65823

A/Accession: S65824

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1275 <DOM>

A/Cross-references: EMBL:M00340; NID:G339767; PIDN:AAA51622.1; PID:G339771

C/Superfamily: pol polyprotein

Query Match 39.5%; Score 672.5; DB 2; Length 1275;
Best Local Similarity 71.4%; Pred. No. 1.6e-44;
Matches 147; Conservative 4; Mismatches 28; Indels 27; Gaps 5;

QY 1 MGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 60
DB 990 VGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 1049
QY 61 ISRIYNELKOIYKKKTNPPIKKWVDMNRHFSKEGIYAAKGMKKYSSSLAIREMOIKTT 120
DB 1050 ISRIYNELKOIYKKKTNPPIKKWVDMNRHFSKEDIYAAKGMKKYSSSLAIREMOIKTT 1109
QY 121 MRVHLTPVRMAIIKKSNNR--RDMDKGNHHSQOTTRTKNOTPHVLTTRW---ILQOS 174
DB 1110 MRVHLTPVRMAIIKKSNNRCMRGCGEIG-----TLHGMWCKLVOP 1152
QY 175 HWYTV---ISDISELMHKTDRIVNLL 197
DB 1153 LMKSVWRFRLDL-ELRIPDPALPLL 1177

RESULT 5

B28096
line-1 protein ORF2 - human

C/Species: Homo sapiens (man)

C/Date: 03-Nov-1988 #sequence_revision 03-Nov-1988 #text_change 30-Sep-1993

C/Accession: B28096

R/Skowronski, J.; Fanning, T.G.; Singer, M.F.

Mol. Cell. Biol. 8, 1385-1397, 1988

A/Title: Unit-length line-1 transcripts in human teratocarcinoma cells.

A/Reference number: A28096; MUID:88246405; PMID:2454389

A/Accession: B28096

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-1275 <SKO>

C/Superfamily: pol polyprotein

Query Match 39.2%; Score 668.5; DB 2; Length 1275;
Best Local Similarity 70.9%; Pred. No. 3.3e-44;
Matches 146; Conservative 4; Mismatches 29; Indels 27; Gaps 5;

QY 1 MGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 60
DB 990 VGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 1049
QY 61 ISRIYNELKOIYKKKTNPPIKKWVDMNRHFSKEGIYAAKGMKKYSSSLAIREMOIKTT 120
DB 1050 ISRIYNELKOIYKKKTNPPIKKWVDMNRHFSKEDIYAAKGMKKYSSSLAIREMOIKTT 1109
QY 121 MRVHLTPVRMAIIKKSNNR--RDMDKGNHHSQOTTRTKNOTPHVLTTRW---ILQOS 174
DB 1110 MRVHLTPVRMAIIKKSNNRCMRGCGEIG-----TLHGMWCKLVOP 1152
QY 175 HWYTV---ISDISELMHKTDRIVNLL 197
DB 1153 LMKSVWRFRLDL-ELRIPDPALPLL 1177

RESULT 6

GNH011

retrovirus-related reverse transcriptase pseudogene - human

C/Species: Homo sapiens (man)

C/Date: 31-Mar-1988 #sequence_revision 04-Jan-1996 #text_change 14-May-1999

C/Accession: A25313

R/Hatori, M.; Kunata, S.; Takenaka, O.; Sakaki, Y.

Nature 321, 625-628, 1986

A/Title: L1 family of repetitive DNA sequences in primates may be derived from a sequence

A:Reference number: A93381; MUID:86230917; PMID:2423883
 A:Accession: A2513
 A:Status: conceptual translation of pseudogene
 A:Molecule type: DNA
 A:Residues: 1-1259 <HAT>
 A:Note: this sequence was constructed from an alignment of published and unpublished seq
 C:Keywords: reverse transcriptase; pseudogene

Query Match 38.6%; Score 658.5; DB 4; Length 1259;
 Best Local Similarity 69.9%; Pred. No. 1.9e-43;

Matches 144; Conservative 6; Mismatches 29; Indels 27; Gaps 5;

QY 1 MGDFMSKTPKAMATKIDKMDILKLSKSPCTAKETIRVNRQLTWEMKIFATYSPDKL 60
 DB MGDFMTPKPKAMATKIDKMDILKLSKSPCTAKETIRVNRQLTWEMKIFATYSPDKL 1048

QY 61 ISRIYVELKQIYKKTNPPIKKVVKDMNRHFSKEGIYAAKGMKKYSSSLAIREMOIKTT 120
 DB 1049 ISRIYVELKQIYKKTNPPIKKVVKDMNRHFSKEDIYAAKGMKKYSSSLAIREMOIKTT 1108

QY 121 MRHILTPVPMALIKKSGNN---RDMDEAGNHSSQOTTRTKNOTPHVLTNRW---ILQOS 174
 DB 1109 MRHILTPVPMALIKKSGNNRDMDEAGNHSSQOTTRTKNOTPHVLTNRW 1151

QY 175 HWVTV---LSDISLHMKHTDRIVNLL 197
 DB 1152 LKSKVNRFLRL-ELRIPFPALPL 1176

RESULT 7

JC7591
 spinal cord-derived growth factor-B precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 24-Aug-2001

C:Accession: J07591
 R:Hamada, T.; Uti-Tel, K.; Imaki, J.; Miyata, Y.
 Biochem. Biophys. Res. Commun. 280, 733-737, 2001

A:Title: Molecular cloning of SCDF-B, a novel growth factor homologous to SCDF/PDGF-C/
 A:Reference number: J07591; MUID:21092670; PMID:11162582

A:Accession: J07591
 A:Molecule type: DNA
 A:Residues: 1-370 <HAM>

A:Cross-references: DDBJ:AB033832
 C:Genetics:

A:Gene: scdgb-B
 F:1-11/Domain: secretory signal sequence #status predicted <SIG>
 F:18-370/Product: spinal cord-derived growth factor-B #status predicted <MAT>
 F:52-170/Region: CUB domain #status predicted

F:272-370/Region: homologous to platelet-derived growth factor/vascular endothelial grow
 F:294-308/Region: conserved motif #status predicted

Query Match 37.4%; Score 637; DB 2; Length 370;
 Best Local Similarity 65.4%; Pred. No. 2.3e-42;
 Matches 125; Conservative 9; Mismatches 11; Indels 46; Gaps 3;

QY 173 QSHWVTVLSDISELMHKT-----DRIVNLMC-----M 200
 DB 180 EIMWESVTSISGVSYNSPVTPPTLIADLDKIAEFTVEDLTKYFNPESQOEDLENN 239

QY 201 YLLT-----VDLRLNDAKRYSCTPRNVNIREELKLANVVEPRCLAV 246
 DB 240 YLLTPRPRGSHRSKVDLRLNDAKRYSCTPRNVNIREELKLANVVEPRCLAV 239

QY 247 QRCGNGCGCTVNRSGCTGSGTKVKKYHEVLOFEPGHIKRRGRAKTALVDIOLDHHER 306
 DB 300 QRCGNGCGCTVNRSGCTGSGTKVKKYHEVLOFEPGHIKRRGRAKTALVDIOLDHHER 359

QY 307 CDCICSSRPR 317
 DB 360 CDCICSSRPR 370

RESULT 8

JC7592

spinal cord-derived growth factor-B precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 24-Aug-2001

C:Accession: J07592
 R:Hamada, T.; Uti-Tel, K.; Imaki, J.; Miyata, Y.
 Biochem. Biophys. Res. Commun. 280, 733-737, 2001

A:Title: Molecular cloning of SCDF-B, a novel growth factor homologous to SCDF/PDGF-C/
 A:Reference number: J07591; MUID:21092670; PMID:11162582

A:Accession: J07592
 A:Molecule type: mRNA
 A:Residues: 1-370 <HAM>
 A:Cross-references: DDBJ:AB052170
 C:Genetics:

A:Gene: scdgb-B
 F:1-11/Domain: secretory signal sequence #status predicted <SIG>
 F:18-370/Product: spinal cord-derived growth factor-B #status predicted <MAT>
 F:52-170/Region: CUB domain #status predicted

F:272-370/Region: homologous to platelet-derived growth factor/vascular endothelial grow
 F:294-308/Region: conserved motif #status predicted

Query Match 34.3%; Score 585; DB 2; Length 370;
 Best Local Similarity 89.4%; Pred. No. 2.7e-38;
 Matches 101; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 205 VDLRLNDAKRYSCTPRNVNIREELKLANVVEPRCLVORCGNGCGCTVNRSGCT 264
 DB 258 VDLRLNDVKKRYSCTPRNVNIREELKLANVVEPRCLVORCGNGCGCTVNRSGCT 317

QY 265 CNSGKTVKKYHEVLOFEPGHIKRRGRAKTALVDIOLDHHERCDICSSRPR 317
 DB 318 CNSGKTVKKYHEVLOFEPGHIKRRGRAKTALVDIOLDHHERCDICSSRPR 370

RESULT 9

GNRL1
 retrovirus-related reverse transcriptase pseudogene - slow loris
 C:Species: Nycticebus coucang (slow loris)
 C:Date: 31-Mar-1988 #sequence_revision 04-Jan-1996 #text_change 14-May-1999

C:Accession: B25313
 R:Hattori, M.; Kuhara, S.; Takenaka, O.; Sakaki, Y.
 Nature 321, 625-628, 1986

A:Title: Ll family of repetitive DNA sequences in primates may be derived from a sequence
 A:Reference number: A93381; MUID:86230917; PMID:2423883

A:Accession: B25313
 A:Status: conceptual translation of pseudogene
 A:Molecule type: DNA

A:Residues: 1-1260 <HAT>
 A:Note: this sequence was constructed from an alignment of six sequences, determined by t
 C:Keywords: reverse transcriptase; pseudogene

Query Match 26.0%; Score 443.5; DB 4; Length 1260;
 Best Local Similarity 50.3%; Pred. No. 1.3e-26;
 Matches 95; Conservative 25; Mismatches 52; Indels 17; Gaps 2;

QY 1 MGDFMSKTPKAMATKIDKMDILKLSKSPCTAKETIRVNRQLTWEMKIFATYSPDKL 60
 DB 989 LGEYFMKRTPOALEAVSKIHVLDILKLSKSPCTAKETIRVNRQLTWEMKIFATYSPDKL 1048

QY 61 ISRIYVELKQIYKKTNPPIKKVVKDMNRHFSKEGIYAAKGMKKYSSSLAIREMOIKTT 120
 DB 1049 ISRIYVELKQIYKKTNPPIKKVVKDMNRHFSKEDIYAAKGMKKYSSSLAIREMOIKTT 1108

QY 121 MRHILTPVPMALIKKSGNNRDMDEAGNHSSQOTTRTKNOTPHVLTNRW---ILQOS 177
 DB 1109 LRYHILTPVPMALIKKSGNNRDMDEAGNHSSQOTTRTKNOTPHVLTNRW 1154

QY 178 TVLSDISEL 186
 DB 1155 DVWRILRLD 1163

RESULT 10

S21348 Probable pol polyprotein-related protein 4 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999

C:Accession: S21348

R:Schultz, E.; Mohr, E.

submitted to the EMBL Data Library, June 1990

A:Reference number: S21345

A:Accession: S21348

A:Molecule type: DNA

A:Residues: 1-275 <SCH>

A:Cross-references: EMBL:X53581; NID:955686; PIDN:CAA37647.1; PID:955690

C:Superfamily: pol polyprotein

C:Keywords: polypeptide

Query Match 25.0%; Score 425.5; DB 2; Length 275;

Best Local Similarity 49.7%; Pred. No. 5.9e-26;

Matches 85; Conservative 28; Mismatches 41; Indels 17; Gaps 2;

12 AMATKATDKMDLTKLSFCTAKETTRVNRQLTEMEKIPATYSFDKGLIRINELKQI 71

Db 2 AYALRSRIDKMDLTKLSFCTAKETTRVNRQLTEMEKIPATYSFDKGLIRINELKQI 61

Qy 72 YKKTKNPKKQVNDNRHFSKEGIYAAKGMKSSSLAIREMOIKTMRVHLTPYMA 131

Db 62 DRRTNNPKKMGSEINKEFTAEBCQMAEKILKCSLIVIREMOIKTMRVHLTPYMA 121

Qy 132 IIKSGNNRMDNAGNHSQOTITRTKNQTPHVLTHRN--IIQSHWTV 179

Db 122 MIKSGNRCWRCGCGERG-----LLHCWMDCKLVQPFKSV 158

RESULT 11

GNMSL

retrovirus-related reverse transcriptase homolog - mouse retrotransposon

N:Alternate names: L1md repetitive element ORF-2; LINE-1 hypothetical protein; ORF 3900;

C:Species: Mus musculus (house mouse)

C:Date: 31-Mar-1989 #sequence_revision 08-Jan-1999 #text_change 16-Jun-2000

C:Accession: B58927; B24906; T49130; A23772; B23430

R:Loeb, D.D.; Padgett, R.W.; Hardies, S.C.; Shehee, W.R.; Comer, M.B.; Edgell, M.H.; Hut

Mol. Cell. Biol. 6, 168-182, 1986

A:Title: The sequence of a large L1md element reveals a tandemly repeated 5' end and sex

A:Reference number: A93072; MUID:87064284; PMID:3023821

A:Accession: B58927

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1281 <LOE1>

A:Cross-references: GB:M13002; NID:9200849; PIDN:AAA6024.1; PID:9804788

A:Note: sequence shown in Fig. 2

R:Martin, S.L.; Martin, S.L.

Gene 153, 261-266, 1995

A:Title: Characterization of a LINE-1 cDNA that originated from RNA present in ribonucle

A:Reference number: 149129; MUID:95180729; PMID:753116

A:Accession: 149130

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-85 'L', 87-358, 'K', 360-706, 'F', 708-735, 'A', 737-760, 'W', 762-927, 'D', 929-1281

A:Cross-references: EMBL:U01647; NID:955806; PIDN:AAA67727.1; PID:9558908

R:Moctez, E.; Kogan, P.K.; Manuelidis, L.

Nucleic Acids Res. 14, 3119-3136, 1986

A:Title: Conservation in the 5' region of the long interspersed mouse L1 repeat: implica

A:Reference number: A23772; MUID:86117689; PMID:3008107

A:Accession: A23772

A:Molecule type: DNA

A:Residues: 'NNQSNHSTNQKEDSHKOR', 1-245, 'K', 247-423, 'STQGNKTKTKTN', 439, 'WDTTRYQS' <MC

A:Cross-references: GB:X0375; NID:952829; PIDN:CAA7363.1; PID:91334115

C:Keywords: reverse transcriptase

Query Match 24.2%; Score 412.5; DB 1; Length 1281;

Best Local Similarity 47.0%; Pred. No. 3.5e-24;

Matches 85; Conservative 30; Mismatches 49; Indels 17; Gaps 2;

2 GKDPMSTPRAMATKAKIDKMDLTKLSFCTAKETTRVNRQLTEMEKIPATYSFDKGLI 61

Db 998 GEKFMNTAAACAVRSIDKMDLTKLSFCTAKETTRVNRQLTEMEKIPATYSFDKGLI 1057

Qy 62 SRIYNEIKQIKKTKNPKKQVNDNRHFSKEGIYAAKGMKSSSLAIREMOIKTMR 121

Db 1058 SNIYKEIKQIKKTKNPKKQVNDNRHFSKEGIYAAKGMKSSSLAIREMOIKTMR 1117

Qy 122 RYHLLTPYMAIIKSGNNRMDNAGNHSQOTITRTKNQTPHVLTHRN--IIQSHWTV 178

Db 1118 RYHLLTPYMAIIKSGNNRMDNAGNHSQOTITRTKNQTPHVLTHRN--IIQSHWTV 1163

Qy 179 V 179

Db 1164 V 1164

RESULT 12

S57662 hypothetical protein 2 - North American opossum (fragment)

C:Species: Didelphis virginiana, Didelphis marsupialis virginiana (North American opossum)

C:Date: 29-Nov-1995 #sequence_revision 01-Mar-1996 #text_change 03-Nov-2000

C:Accession: S57662

R:Doerner, M.; Paabo, S.

submitted to the EMBL Data Library, April 1995

A:Description: Nucleotide sequence of a Marsupial LINE-1 element and the evolution of pl

A:Reference number: S57662

A:Accession: S57662

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-392 <DOE>

A:Cross-references: EMBL:Z48955; NID:9897720; PID:9897721

C:Genetics:

A:Insertions: 183/3

C:Superfamily: pol polyprotein

Query Match 21.9%; Score 374; DB 2; Length 392;

Best Local Similarity 55.4%; Pred. No. 9.5e-22;

Matches 72; Conservative 21; Mismatches 37; Indels 0; Gaps 0;

2 GKDPMSTPRAMATKAKIDKMDLTKLSFCTAKETTRVNRQLTEMEKIPATYSFDKGLI 61

Db 113 GKDPMSTPRAMATKAKIDKMDLTKLSFCTAKETTRVNRQLTEMEKIPATYSFDKGLI 172

Qy 62 SRIYNEIKQIKKTKNPKKQVNDNRHFSKEGIYAAKGMKSSSLAIREMOIKTMR 121

Db 173 TQYKEIKQIKKTKNPKKQVNDNRHFSKEGIYAAKGMKSSSLAIREMOIKTMR 232

Qy 122 RYHLLTPYMA 131

Db 233 RYHLLTPYMA 242

RESULT 13

F41925 hypothetical protein 2 - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 26-Aug-1999

C:Accession: F41925

R:Dotzlaw, H.; Alkhalaf, M.; Murphy, L.C.

Mol. Endocrinol. 6, 773-785, 1992

A:Title: Characterization of estrogen receptor variant mRNAs from human breast cancers.

A:Reference number: A41925; MUID:92293154; PMID:1603086

A:Accession: F41925

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-92 <DOT>

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 12:32:13 ; Search time 70 Seconds

(without alignments)
603,435 Million cell updates/sec

Title: US-10-083-853B-1

Perfect score: 1704
Sequence: 1 MGKDFMSKTPKAWATAKID.....DIQLDHERDCICSSRPPR 317

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	818.5	48.0	939	22	ABG12508 Novel human diagno
2	818.5	48.0	939	22	ABG14439 Novel human diagno
3	818.5	48.0	940	22	ABG07054 Novel human diagno
4	818.5	48.0	736	22	ABG04462 Novel human diagno
5	812.5	47.7	969	22	ABG14779 Novel human diagno
6	812.5	47.7	1074	22	ABG10255 Novel human diagno
7	804.5	47.2	1074	22	ABG07890 Novel human diagno
8	804	47.1	853	22	ABG17719 Novel human diagno
9	803	46.0	22	22	ABG14765 Novel human diagno
10	783.5	46.0	1680	22	ABG17645 Novel human diagno

11	783.5	46.0	1825	22	ABG11906 Novel human diagno
12	779	45.7	279	22	ABG15456 Novel human diagno
13	771	45.2	827	22	AA130932 Novel human diagno
14	760.5	44.6	1067	22	ABG18495 Novel human diagno
15	758.5	44.5	1356	22	ABG10381 Novel human diagno
16	758.5	44.5	1721	22	ABG08687 Novel human diagno
17	757.5	44.5	1655	22	ABG14955 Novel human diagno
18	755.5	44.3	536	22	ABG07412 Novel human diagno
19	755.5	44.3	536	22	ABG12326 Novel human diagno
20	755.5	44.3	1361	22	ABG23198 Novel human diagno
21	755.5	44.3	2764	22	ABG10129 Novel human diagno
22	755	43.7	1261	22	ABG07064 Novel human diagno
23	744.5	43.7	1565	22	ABG06288 Novel human diagno
24	744	43.7	1669	22	ABG13314 Novel human diagno
25	736	43.2	894	22	ABG08978 Novel human diagno
26	736	43.2	894	22	ABG13059 Novel human diagno
27	730	42.8	662	22	ABG05640 Novel human diagno
28	730	42.8	662	22	ABG07422 Novel human diagno
29	730	42.8	662	22	ABG14256 Novel human diagno
30	728	42.7	944	22	ABG02301 Novel human diagno
31	728	42.7	944	22	ABG18485 Novel human diagno
32	727	42.7	671	22	ABG17890 Novel human diagno
33	726	42.6	908	22	ABG05638 Novel human diagno
34	726	42.6	908	22	ABG12511 Novel human diagno
35	725	42.5	597	22	ABG07850 Novel human diagno
36	725	42.5	597	22	ABG10065 Novel human diagno
37	725	42.5	703	22	ABG14431 Novel human diagno
38	725	42.5	703	22	ABG23675 Novel human diagno
39	725	42.5	754	22	ABG12695 Novel human diagno
40	724	42.5	452	22	ABG06567 Novel human diagno
41	724	42.5	959	22	ABG02299 Novel human diagno
42	720	42.3	2678	22	ABG11124 Novel human diagno
43	718.5	42.2	1344	22	ABG13058 Novel human diagno
44	718	42.1	775	22	ABG02472 Novel human diagno
45	715	42.0	1078	22	ABG17206 Novel human diagno

ALIGNMENTS

RESULT 1
ABG12508
ID ABG12508 standard; Protein; 939 AA.
XX
AC ABG12508;
XX
DT 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #12499.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
PR
XX (HYSE-) HYSEQ INC.
XX
XX Dmanac RT, Liu C, Tang YF,
PI
XX WPI, 2001-639362/73.
XX
XX N-PSDB; AAST6695.
DR
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 42867; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridization probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 939 AA;
Query Match 48.0%; Score 818.5; DB 22; Length 939;
Best Local Similarity 82.4%; Pred. No. 2.1e-75;
Matches 159; Conservative 9; Mismatches 18; Indels 7; Gaps 1;
QY 1 MGKPMKTPKAMATKAKIDKMDLILKSFCTAKETTRVNRQUTLWEMKIFATYSPDKGL 60
DB 685 MGKPMKTPKAMATKAKIDKMDLILKSFCTAKETTRVNRQUTLWEMKIFATYSPDKGL 744
QY 61 ISRIYNELKQIYKKKTNPPIKKNVWKNRHRFSKGIYAAKGMKKYSSSLAIREMOIKTT 120
DB 745 ISRIHNEIKQIYKKKTNPPIKKNVWKNRHRFSKGIYAAKGMKKYSSSLAIREMOIKTT 804
QY 121 MRHYLTPVRMAIIRKSGNNRDMDEAGNHSQOITTRTKNQTPHVLTNRWL-----QQ 173
DB 805 MRHYLTPVRMAVIRKSGNNRDMDEAGNHSQOITTRTKNQTPHVLTNRWLNEITWQ 864
QY 174 SHWTVLSDISEL 186
DB 865 GEYHTLGTIVSEL 877
RESULT 2
ABG14439
ID ABG14439 standard; Protein; 939 AA.
XX
AC ABG14439;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #14430.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.

XX
PA (HYSE-) HYSEQ INC.
XX
PT Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS78626.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 44798; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridization probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 939 AA;
Query Match 48.0%; Score 818.5; DB 22; Length 939;
Best Local Similarity 82.4%; Pred. No. 2.1e-75;
Matches 159; Conservative 9; Mismatches 18; Indels 7; Gaps 1;
QY 1 MGKPMKTPKAMATKAKIDKMDLILKSFCTAKETTRVNRQUTLWEMKIFATYSPDKGL 60
DB 685 MGKPMKTPKAMATKAKIDKMDLILKSFCTAKETTRVNRQUTLWEMKIFATYSPDKGL 744
QY 61 ISRIYNELKQIYKKKTNPPIKKNVWKNRHRFSKGIYAAKGMKKYSSSLAIREMOIKTT 120
DB 745 ISRIHNEIKQIYKKKTNPPIKKNVWKNRHRFSKGIYAAKGMKKYSSSLAIREMOIKTT 804
QY 121 MRHYLTPVRMAIIRKSGNNRDMDEAGNHSQOITTRTKNQTPHVLTNRWL-----QQ 173
DB 805 MRHYLTPVRMAVIRKSGNNRDMDEAGNHSQOITTRTKNQTPHVLTNRWLNEITWQ 864
QY 174 SHWTVLSDISEL 186
DB 865 GEYHTLGTIVSEL 877
RESULT 3
ABG07054
ID ABG07054 standard; Protein; 940 AA.
XX
AC ABG07054;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #7045.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX

OS Homo sapiens.
 XX WO200175067-A2.
 PN 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT,
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS71241.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 Claim 20; SEQ ID No 37413; 103bp; English.
 The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensic, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 940 AA;
 SQ
 Query Match 48.0%; Score 818.5; DB 22; Length 940;
 Best Local Similarity 82.4%; Pred. No. 2.1e-75;
 Matches 159; Conservative 9; Mismatches 18; Indels 7; Gaps 1;
 Oy 1 MGKDFMSKTPKAMATRAKIDKMDLILKLSFCTAKETTRIVNRQLTWEKIFATYSPDKGL 60
 Db 686 MGKDFMSKTPKAMATRAKIDKMDLILKLSFCTAKETTRIVNRQLTWEKIFATYSPDKGL 745
 Oy 61 ISRIYNELKQIYKKTKNPIKQVVKDMNRRHFSKEGIYAAXKMKYSSSLAREMOIKTT 120
 Db 746 ISRIYNELKQIYKKTKNPIKQVVKDMNRRHFSKEGIYAAXKMKYSSSLAREMOIKTT 805
 Oy 121 MRYHLTPVMAIILKSGNNRDMDEAGNHSQOTITRTKQTPHVLTHRWIL-----QQ 173
 Db 806 MRYHLTPVMAIILKSGNNRDMDEAGNHSQOTITRTKQTPHVLTHRWELNNEITWTQE 865
 Oy 174 SHWVTVLSDISEL 186
 Db 866 GEYHTLGTVTSSEL 878
 RESULT 4
 ABG04462
 ID ABG04462 standard; Protein; 736 AA.

XX ABG04462;
 AC 13-FEB-2002 (first entry)
 XX Novel human diagnostic protein #4453.
 DT
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS WO200175067-A2.
 PN 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT,
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS68649.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 Claim 20; SEQ ID No 34821; 103bp; English.
 The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensic, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX Sequence 736 AA;
 SQ
 Query Match 48.0%; Score 818; DB 22; Length 736;
 Best Local Similarity 82.1%; Pred. No. 1.7e-75;
 Matches 160; Conservative 7; Mismatches 18; Indels 10; Gaps 2;
 Oy 1 MGKDFMSKTPKAMATRAKIDKMDLILKLSFCTAKETTRIVNRQLTWEKIFATYSPDKGL 60
 Db 176 MGKDFMSKTPKAMATRAKIDKMDLILKLSFCTAKETTRIVNRQLTWEKIFATYSPDKGL 235
 Oy 61 ISRIYNELKQIYKKTKNPIKQVVKDMNRRHFSKEGIYAAXKMKYSSSLAREMOIKTT 120
 Db 236 ISRIYNELKQIYKKTKNPIKQVVKDMNRRHFSKEGIYAAXKMKYSSSLAREMOIKTT 295
 Oy 121 MRYHLTPVMAIILKSGNNRDMDEAGNHSQOTITRTKQTPHVLTHRWIL--QSHWV-- 177

Db 296 MRYHLTPVRMAIIKSGNNRDMDEGNHSGQTTARTKQTLPHVLTTHRWELNENWTWQE 355
Qy 178 -----TVLSDISF 185
Db 356 GEHHTLETFVGDMD 370

RESULT 5

ABG14779
ID ABG14779 standard; Protein; 969 AA.
AC ABG14779;
XX
XX
XX 18-FEB-2002 (first entry)
DT
DE Novel human diagnostic protein #14770.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX MO200175067-A2.
PN
PD 11-OCT-2001.
XX

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Dermanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.
DR N-PSDB; AAS78966.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID NO 45138; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 969 AA;

Query Match 47.7%; Score 812.5; DB 22; Length 969;
Best Local Similarity 88.7%; Pred. No. 9.1e-75;
Matches 157; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

Qy 1 MGKDFMSKTPKAMATKXIDKMDLIKLSFCTAKETTRIVNROLTEWIKIFATYSPDKGL 60
Db 779 MGKDFMSKTPKAMATKXIDKMDLIKLSFCTAKETTRIVNROLTEWIKIFATYSPDKGL 838
Qy 61 ISRIYNELKQIYKKTKNPIKXWVKDMNRHFSKEGIYAARKHKKYSSSLAIREMOIKTT 120
Db 839 ISRIYNELKQIYKKTKNPIKXWVKDMNRHFSKEDIYAAARKHKKCSSSLAIREMOIKTT 898
Qy 121 MRYHLTPVRMAIIKSGNNRDMDEGNHSGQTTARTKQTLPHVLTTHRWELNENWTWQE 176
Db 899 MRYHLTPVRMAIIKSGNNRDMDEGNHSGQTTARTKQTLPHVLTTHRWELNENWTW 955

RESULT 6

ABG10255
ID ABG10255 standard; Protein; 1074 AA.

AC ABG10255;
XX
XX
XX 13-FEB-2002 (first entry)
DT
DE Novel human diagnostic protein #10246.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX MO200175067-A2.
PN
PD 11-OCT-2001.
XX

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Dermanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.
DR N-PSDB; AAS74442.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID NO 40614; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 1074 AA;
Query Match 47.7%; Score 812.5; DB 22; Length 1074;
Best Local Similarity 82.9%; Pred. No. 1.1e-74;
Matches 160; Conservative 8; Mismatches 18; Indels 7; Gaps 1;
QY 1 MGDFMSKTPKAMATKAKIDKMDLKLKSFCTAKETITIRVNRQLTWEKIFATYSPDKGL 60
DB MGDFMSKTPKAMATKAKIDKMDLKLKSFCTAKETITIRVNRQLTWEKIFATYSPDKGL 879
QY 61 ISRIYNELKQIYKKTKNPIKKVVDNMNRHPSKEGIYAAKGMKKYSSSLAIRMOIKTT 120
DB ISRIYNELKQIYKKTKNPIKKVVDNMNRHPSKEGIYAAKGMKKYSSSLAIRMOIKTT 939
QY 121 MRYHLTPVPMALIKKSGNNRDMDEAGNHSOQTITRTKQTPHYLTHRWTL-----QQ 173
DB 940 VRHLLTPVPMALIKKSGNNRDMDEAGNHSOQTITRTKQTPHYLTHRWTLNEITWTQ 999
QY 174 SHWTVLSDISEL 186
DB 1000 GEYHTLTGTVTSEL 1012
RESULT 7
ABG07890
ID ABG07890 standard; Protein; 1074 AA.
XX ABG07890;
AC
XX
XX 13-FEB-2002 (first entry)
DT
XX
XX Novel human diagnostic protein #7881.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM Food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX
XX MO200175067-A2.
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX
XX WPI, 2001-639362/73.
DR N-PSDB; AAS72077.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
PS Claim 20; SEQ ID No 38249; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (II) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcr_sequences.
XX
SQ Sequence 1074 AA;
Query Match 47.2%; Score 804.5; DB 22; Length 1074;
Best Local Similarity 81.9%; Pred. No. 7.1e-74;
Matches 158; Conservative 9; Mismatches 19; Indels 7; Gaps 1;
QY 1 MGDFMSKTPKAMATKAKIDKMDLKLKSFCTAKETITIRVNRQLTWEKIFATYSPDKGL 60
DB MGDFMSKTPKAMATKAKIDKMDLKLKSFCTAKETITIRVNRQLTWEKIFATYSPDKGL 879
QY 61 ISRIYNELKQIYKKTKNPIKKVVDNMNRHPSKEGIYAAKGMKKYSSSLAIRMOIKTT 120
DB ISRIYNELKQIYKKTKNPIKKVVDNMNRHPSKEGIYAAKGMKKYSSSLAIRMOIKTT 939
QY 121 MRYHLTPVPMALIKKSGNNRDMDEAGNHSOQTITRTKQTPHYLTHRWTL-----QQ 173
DB 940 VRHLLTPVPMALIKKSGNNRDMDEAGNHSOQTITRTKQTPHYLTHRWTLNEITWTQ 999
QY 174 SHWTVLSDISEL 186
DB 1000 GEYHTLTGTVTSEL 1012
RESULT 8
ABG17719
ID ABG17719 standard; Protein; 482 AA.
XX ABG17719;
AC
XX
XX 18-FEB-2002 (first entry)
DT
XX
XX Novel human diagnostic protein #17710.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM Food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX
XX MO200175067-A2.
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX
XX WPI, 2001-639362/73.
DR N-PSDB; AAS81906.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
PS Claim 20; SEQ ID No 48078; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 482 AA;

Query Match 47.2%; Score 804; DB 22; Length 482;
Best Local Similarity 91.1%; Pred. No. 2.6e-74;
Matches 154; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MGKDFMSKTPKAMATKAKIDKMDLILKLSFCTAKETTRVNRDLTEWEKIFATYSPDKGL 60
DB 237 MGKDFMSKTPKAMATKAKIDKMDLILKLSFCTAKETTRVNRDLTEWEKIFATYSPDKGL 296
QY 61 ISRIYNELKOIYKKKTNPPIKKNVNDNRHFSKEGIYAAKGMKKYSSLAIREMOIKTT 120
DB 297 ISRIYNELKOIYKKKTNPPIKKNVNDNRHFSKEDIYAAKGMKKYSSLAIREMOIKTT 356
QY 121 MRYHLTPVRMAIYKSGNNDMDAGNHSQOITTRTKNQTPIHLTHW 169
DB 357 MRYHLTPVRMAIYKSGNNDMDAGNHSQOITTRTKNQTPIHLTHW 405

RESULT 9
ABG14765
ID ABG14765 standard; Protein: 853 AA.

XX ABG14765;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #14756.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

PN MO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

XX N-PSDB; AAS76952.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity
XX Claim 20; SEQ ID No 45124; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probe,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 853 AA;

Query Match 47.1%; Score 803; DB 22; Length 853;
Best Local Similarity 91.0%; Pred. No. 7.3e-74;
Matches 152; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MGKDFMSKTPKAMATKAKIDKMDLILKLSFCTAKETTRVNRDLTEWEKIFATYSPDKGL 60
DB 686 MGKDFMSKTPKAMATKAKIDKMDLILKLSFCTAKETTRVNRDLTEWEKIFATYSPDKGL 745
QY 61 ISRIYNELKOIYKKKTNPPIKKNVNDNRHFSKEGIYAAKGMKKYSSLAIREMOIKTT 120
DB 746 ISRIYNELKOIYKKKTNPPIKKNVNDNRHFSKEDIYAAKGMKKYSSLAIREMOIKTT 805
QY 121 MRYHLTPVRMAIYKSGNNDMDAGNHSQOITTRTKNQTPIHLTHW 167
DB 806 MRYHLTPVRMAIYKSGNNDMDAGNHSQOITTRTKNQTPIHLTHW 852

RESULT 10
ABG17645
ID ABG17645 standard; Protein: 1680 AA.

XX ABG17645;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #17636.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

PN MO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR MPI; 2001-639362/73.
DR N-PSDB; AAS81832.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
PS Claim 20; SEQ ID No 48004; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (II) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue; as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1680 AA;
Query Match 46.0%; Score 783.5; DB 22; Length 1680;
Best Local Similarity 86.9%; Pred. No. 2e-71;
Matches 153; Conservative 5; Mismatches 17; Indels 1; Gaps 1;
QY 1 MGKDFMSKTPKAMATYAKIDKMDLILKLSFCTAKETTRVNRQLTWEMKIFATYSDKGL 60
DB 990 MGKDFISKTPKAMATYAKIDKMDLILKLSFCTAKETTRVNRQPTWEMKIFATYSDKGL 1049
QY 61 ISRIYVELKQIYKKKTKN-PIKCVKDMNRHPSKEGIVYAKKMKYSSLAIREMOIKT 119
DB 1050 ISRIYVELKQIYKKKTKNPIKCVKDMNRHPSKEGIVYAKKMKYSSLAIREMOIKT 1109
QY 120 TMRVHLTPVRMAIIKKSGNNRDMDEAGNHSQOTITRTKNQTPHVLTHRMILQOSH 175
DB 1110 TMRVHLTPVRMAIIKKSGNNRDMDEAGNHSQOTITRTKNQTPHVLTHRMILQOSH 1165
RESULT 11
ABG11906
ID ABG11906 standard; Protein; 1825 AA.
AC ABG11906;
XX
XX 18-FEB-2002 (first entry)
DE Novel human diagnostic protein #11897.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX
XX

PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YR;
XX
XX MPI; 2001-639362/73.
DR N-PSDB; AAS76093.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
PS Claim 20; SEQ ID No 42265; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue; as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1825 AA;
Query Match 46.0%; Score 783.5; DB 22; Length 1825;
Best Local Similarity 85.9%; Pred. No. 2.3e-71;
Matches 152; Conservative 5; Mismatches 19; Indels 1; Gaps 1;
QY 1 MGKDFMSKTPKAMATYAKIDKMDLILKLSFCTAKETTRVNRQLTWEMKIFATYSDKGL 60
DB 426 MGKDFMSKTPKAMATYAKIDKMDLILKLSFCTAKETTRVNRQPTWEMKIFATYSDKGR 485
QY 61 ISRIYVELKQIYKKKTKNPIKCVKDMNRHPSKEGIVYAKKMKYSSLAIREMOIKT 120
DB 486 ISRIYVELKQIYKKKTKNPIKCVKDMNRHPSKEGIVYAKKMKYSSLAIREMOIKT 545
QY 121 KRVYHLTPVRMAIIKKSGNNRDMDEAGNHSQOTITRTKNQTPHVLTHRMILQOSH-W 176
DB 546 KRVYHLTPVRMAIIKKSGNNRDMDEAGNHSQOTITRTKNQTPHVLTHRMILQOSH-W 602
RESULT 12
ABG15456
ID ABG15456 standard; Protein; 279 AA.
AC ABG15456;
XX
XX 13-FEB-2002 (first entry)
DE Novel human diagnostic protein #15447.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX

Qy	61	ISRIYNELKOIYKKKTNP	IKKWVNDNBRHPSKEGIYA	AKKMKYSSLAIREMOIKTT	120
Db	727	ISRIYNELKOIYKKKTNP	IKKWVNDNBRHPSKEDIYA	AKKMKYSSLAIREMOIKTT	786
Qy	121	MRVHLTPVRMAIIKKS	GNNRMDMDEAGNHSSQOIT	TRTKNOTPHVLT	HRMILQOSH-W 176
Db	787	MRVHLTPVRMAIIKKS	GNNRMDMDEAGNHSSQOIT	TRTKNOTPHVLT	HRMILQOSH-W 843

Search completed: June 5, 2003, 12:41:51
 Job time : 72 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 12:40:39 , Search time 26 Seconds
(without alignments)
358.733 Million cell updates/sec

Title: US-10-083-853b-1

Sequence: 1 MCKDFMSKPKPMATKAKID.....DIQLDHERDCICSSRPPR 317

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database :
- 1: Issued Patents_AA:*
 - 2: /cgn2_6/ptodata/1/aa/5A_COMB.pep:*
 - 3: /cgn2_6/ptodata/1/aa/5B_COMB.pep:*
 - 4: /cgn2_6/ptodata/1/aa/6A_COMB.pep:*
 - 5: /cgn2_6/ptodata/1/aa/6B_COMB.pep:*
 - 6: /cgn2_6/ptodata/1/aa/PTUS_COMB.pep:*
 - 6: /cgn2_6/ptodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	637	37.4	370	4	US-09-457-066-37 Sequence 37, Appl
2	637	37.4	370	4	US-09-540-224-2 Sequence 2, Appl
3	588	34.5	370	4	US-09-540-224-4 Sequence 4, Appl
4	318.5	18.7	345	4	US-09-040-220-2 Sequence 2, Appl
5	318.5	18.7	345	4	US-09-457-066-2 Sequence 2, Appl
6	318.5	18.7	345	4	US-09-265-686-2 Sequence 2, Appl
7	318.5	18.7	345	4	US-09-540-224-5 Sequence 5, Appl
8	314.5	18.5	345	4	US-09-457-066-43 Sequence 43, Appl
9	118	6.9	24	4	US-09-540-224-9 Sequence 9, Appl
10	114.5	6.7	354	4	US-08-915-795-5 Sequence 5, Appl
11	112.5	6.6	325	4	US-08-915-795-3 Sequence 3, Appl
12	106.5	6.2	321	4	US-08-915-795-9 Sequence 9, Appl
13	106.5	6.2	358	4	US-08-915-795-8 Sequence 8, Appl
14	102	6.0	109	4	US-09-469-186-1 Sequence 1, Appl
15	96.5	5.7	350	2	US-08-999-811-4 Sequence 4, Appl
16	96.5	5.7	350	2	US-08-824-996-2 Sequence 2, Appl
17	96.5	5.7	350	3	US-09-042-105-4 Sequence 4, Appl
18	96.5	5.7	350	4	US-08-510-133A-33 Sequence 33, Appl
19	96.5	5.7	350	4	US-08-510-133A-33 Sequence 33, Appl
20	96.5	5.7	350	4	US-08-510-133A-33 Sequence 33, Appl
21	96.5	5.7	419	2	US-08-999-811-2 Sequence 2, Appl
22	96.5	5.7	419	3	US-09-042-105-2 Sequence 2, Appl
23	96.5	5.7	419	3	US-09-042-105-18 Sequence 18, Appl
24	96.5	5.7	419	4	US-08-795-430-8 Sequence 8, Appl
25	96.5	5.7	419	4	US-08-510-133A-35 Sequence 35, Appl
26	96.5	5.7	419	4	US-08-510-133A-35 Sequence 35, Appl
27	96.5	5.7	419	5	US-08-601-132-33 Sequence 33, Appl
					PCT-US96-09001-2 Sequence 2, Appl

28	95.5	5.6	102	1	US-08-469-427A-2 Sequence 2, Appl
29	95.5	5.6	102	2	US-08-609-443B-2 Sequence 2, Appl
30	95.5	5.6	102	2	US-08-569-063C-2 Sequence 2, Appl
31	95.5	5.6	102	4	US-08-851-896-2 Sequence 2, Appl
32	95.5	5.6	133	1	US-08-469-427A-9 Sequence 9, Appl
33	95.5	5.6	133	2	US-08-609-443B-9 Sequence 9, Appl
34	95.5	5.6	133	2	US-08-569-063C-9 Sequence 9, Appl
35	95.5	5.6	133	4	US-08-851-896-9 Sequence 9, Appl
36	95.5	5.6	188	1	US-08-469-427A-5 Sequence 5, Appl
37	95.5	5.6	188	2	US-08-609-443B-5 Sequence 5, Appl
38	95.5	5.6	188	2	US-08-569-063C-5 Sequence 5, Appl
39	95.5	5.6	188	4	US-08-851-896-5 Sequence 5, Appl
40	95.5	5.6	207	2	US-08-609-443B-13 Sequence 13, Appl
41	95.5	5.6	207	2	US-08-569-063C-13 Sequence 13, Appl
42	95.5	5.6	207	4	US-08-851-896-13 Sequence 13, Appl
43	94	5.5	195	1	US-08-469-427A-7 Sequence 7, Appl
44	94	5.5	195	2	US-08-609-443B-7 Sequence 7, Appl
45	94	5.5	195	2	US-08-569-063C-7 Sequence 7, Appl

ALIGNMENTS

```
RESULT 1
US-09-457-066-37
; Sequence 37, Application US/09457066
; Patent No. 643673
; GENERAL INFORMATION:
; APPLICANT: Gao, Zeren
; APPLICANT: Hart, Charles E.
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Gilbertson, Debra G.
; APPLICANT: West, James W.
; TITLE OF INVENTION: GROWTH FACTOR HOMOLOG ZVEGF3
; FILE REFERENCE: 98-60
; CURRENT APPLICATION NUMBER: US/09/457,066
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-457-066-37

Query Match
Best Local Similarity 37.4%; Score 637; DB 4; Length 370;
Matches 125; Conservative 9; Mismatches 11; Indels 46; Gaps 3;

QY 173 QSHWTVLSDISLHMKT-----DRIYVLMC-----M 200
Db 180 ETWMEVTSISISGVSYSPVTDPTLIADALDKIAEFTVVEDLTKYFNPESQOEDLENN 239
QY 201 YLT-----VDLRINDAKRYSCPTPNYSVNIIEELKIANVFFPRLTIV 246
Db 240 YLTPTPRGRSYDRKSKVDLRINDAKRYSCPTPNYSVNIIEELKIANVFFPRLTIV 299
QY 247 QRCGNGCGCTVWNRCTCNSGKTIVKYEHLVLFEPGHIKRRGRATMALVDIQLDHER 306
Db 300 QRCGNGCGCTVWNRCTCNSGKTIVKYEHLVLFEPGHIKRRGRATMALVDIQLDHER 359
QY 307 CDCICSSRPPR 317
Db 360 CDCICSSRPPR 370

RESULT 2
US-09-540-224-2
; Sequence 2, Application US/09540224
; Patent No. 6468543
; GENERAL INFORMATION:
```

APPLICANT: Gilbertson, Debra G.
TITLE OF INVENTION: METHODS FOR PROMOTING GROWTH OF BONE,
TITLE OF INVENTION: LIGAMENT AND CARTILAGE USING ZVEGF4
FILE REFERENCE: 00-28
CURRENT APPLICATION NUMBER: US/09/540,224
CURRENT FILING DATE: 2000-03-31
EARLIER APPLICATION NUMBER: US 60/180,169
EARLIER FILING DATE: 2000-02-04
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 370
TYPE: PRT
ORGANISM: Homo sapiens
US-09-540-224-2

Query Match 37.4%; Score 637; DB 4; Length 370;
Best Local Similarity 65.4%; Pred. No. 5.8e-59;
Matches 125; Conservative 9; Mismatches 11; Indels 46; Gaps 3;

Qy 173 QSHWTVLSDISELMHKT-----DRVNLIMC-----M 200
Db 180 ETNWSYTSISGVSNPSVYDPTLADALDKIAEPDVEDLKYENPESWQEDLENN 239
Qy 201 YLLT-----VDLRLNDADKRYSCPTPNYSVNIREEKLANVFFPRCLLV 246
Db 240 YLDPRIRKGRSYHDKRSKYVDLRLNDADKRYSCPTPNYSVNIREEKLANVFFPRCLLV 299
Qy 247 QRCGNGCGGTVMNRSCCTCNSGKTVKYYEVLQEPFGHIKRGRAKTALVDIOLDHHER 306
Db 300 QRCGNGCGGTVMNRSCCTCNSGKTVKYYEVLQEPFGHIKRGRAKTALVDIOLDHHER 359
Qy 307 CDCICSSRPPR 317
Db 360 CDCICSSRPPR 370

RESULT 3
US-09-540-224-4
Sequence 4, Application US/09540224
Patent No. 6468543
GENERAL INFORMATION:
APPLICANT: Gilbertson, Debra G.
TITLE OF INVENTION: METHODS FOR PROMOTING GROWTH OF BONE,
TITLE OF INVENTION: LIGAMENT AND CARTILAGE USING ZVEGF4
FILE REFERENCE: 00-28
CURRENT APPLICATION NUMBER: US/09/540,224
CURRENT FILING DATE: 2000-03-31
EARLIER APPLICATION NUMBER: US 60/180,169
EARLIER FILING DATE: 2000-02-04
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 370
TYPE: PRT
ORGANISM: Mus musculus
US-09-540-224-4

Query Match 34.5%; Score 588; DB 4; Length 370;
Best Local Similarity 90.3%; Pred. No. 8.6e-54;
Matches 102; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
Qy 205 VDLRLNDADKRYSCPTPNYSVNIREEKLANVFFPRCLLVORCGNGCGGTVMNRSC 264
Db 258 VDLRLNDADKRYSCPTPNYSVNIREEKLANVFFPRCLLVORCGNGCGGTVMNRSC 317
Qy 265 CNSGKTVKYYEVLQEPFGHIKRGRAKTALVDIOLDHHERCDIC 317
Db 318 CNSGKTVKYYEVLQEPFGHIKRGRAKTALVDIOLDHHERCDICSSRPPR 370

RESULT 4
US-09-040-220D-2
Sequence 2, Application US/09040220D
Patent No. 691311
GENERAL INFORMATION:
APPLICANT: Ferrara, Napoleone
TITLE OF INVENTION: NOVEL POLYPEPTIDES HAVING HOMOLOGY TO VASCULAR
TITLE OF INVENTION: ENDOTHELIAL CELL GROWTH FACTOR AND BONE MORPHOGENETIC
TITLE OF INVENTION: PROTEIN 1 AND NUCLEIC ACIDS ENCODING SAME, THEIR USES,
FILE REFERENCE: P1122
CURRENT APPLICATION NUMBER: US/09/040,220D
CURRENT FILING DATE: 1998-03-17
NUMBER OF SEQ ID NOS: 8
SEQ ID NO 2
LENGTH: 345
TYPE: PRT
ORGANISM: Human
US-09-040-220D-2

Query Match 18.7%; Score 318.5; DB 4; Length 345;
Best Local Similarity 53.3%; Pred. No. 2.2e-25;
Matches 57; Conservative 14; Mismatches 33; Indels 3; Gaps 1;

Qy 205 VDLRLNDADKRYSCPTPNYSVNIREEKLANVFFPRCLLVORCGNGCGGTVMNRSC 264
Db 236 VDLRLNLEEVRLYSCTPNYSVNIREEKRTDTTFWGCCLLVKRCGNCACCLHNCBQ 295
Qy 265 CNSGKTVKYYEVLQEPFGHIKRGRAKTALVDIOLDHHERCDIC 311
Db 296 CVPEKTVKYYEVLQELRP---KTGVRLGHLSTVDVALEHHEBECDCVC 339

RESULT 5
US-09-457-066-2
Sequence 2, Application US/09457066
Patent No. 6432673
GENERAL INFORMATION:
APPLICANT: Gao, Zeren
APPLICANT: Hart, Charles E.
APPLICANT: Piddington, Christopher S.
APPLICANT: Sheppard, Paul O.
APPLICANT: Shoemaker, Kimberly E.
APPLICANT: Gilbertson, Debra G.
TITLE OF INVENTION: GROWTH FACTOR HOMOLOGY ZVEGF3
FILE REFERENCE: 98-60
CURRENT APPLICATION NUMBER: US/09/457,066
CURRENT FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 345
TYPE: PRT
ORGANISM: Homo sapiens
US-09-457-066-2

Query Match 18.7%; Score 318.5; DB 4; Length 345;
Best Local Similarity 53.3%; Pred. No. 2.2e-25;
Matches 57; Conservative 14; Mismatches 33; Indels 3; Gaps 1;

Qy 205 VDLRLNDADKRYSCPTPNYSVNIREEKLANVFFPRCLLVORCGNGCGGTVMNRSC 264
Db 236 VDLRLNLEEVRLYSCTPNYSVNIREEKRTDTTFWGCCLLVKRCGNCACCLHNCBQ 295
Qy 265 CNSGKTVKYYEVLQEPFGHIKRGRAKTALVDIOLDHHERCDIC 311
Db 296 CVPEKTVKYYEVLQELRP---KTGVRLGHLSTVDVALEHHEBECDCVC 339

RESULT 6
US-09-265-686-2

```
; Sequence 2, Application us/09265686
; Patent No. 6455283
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Napoleone
; APPLICANT: Kuo, Sophia S.
; TITLE OF INVENTION: POLYPEPTIDES HOMOLOGOUS TO VEGF AND BMP1
; FILE REFERENCE: P1122P2
; CURRENT APPLICATION NUMBER: US/09/265,686
; CURRENT FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: US 09/040,220
; PRIOR FILING DATE: 1998-03-17
; PRIOR APPLICATION NUMBER: US 09/184,216
; PRIOR FILING DATE: 1998-11-02
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 2
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Human
US-09-265-686-2
```

```
Query Match 18.7%; Score 318.5; DB 4; Length 345;
Best Local Similarity 53.3%; Pred. No. 2.2e-25;
Matches 57; Conservative 14; Mismatches 33; Indels 3; Gaps 1;

Qy 205 VDLRLNDKARYSCPTPRNSVNIREEIKLANVFPFRCILVORCGNCGCTVNNRSC 264
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 236 VDLNLTEBRALYSCPTPRNSVSIREBLKRTDTITFWGCLLVKRCGNCACCLHNCNEQ 295

Qy 265 CNSGKTVKKYHEVLOFEFGHIKRGAKTMAVDIOLDHHERDCIC 311
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 296 CVPSTKYTKYHEVLOLRP---KTGVKGLHKSITDVALHEHBECDVC 339
```

```
RESULT 7
US-09-540-224-5
; Sequence 5, Application us/09540224
; Patent No. 6468543
; GENERAL INFORMATION:
; APPLICANT: Gilbertson, Debra G.
; APPLICANT: Hart, Charles E.
; TITLE OF INVENTION: METHODS FOR PROMOTING GROWTH OF BONE,
; TITLE OF INVENTION: LIGAMENT AND CARTILAGE USING ZVEGF4
; FILE REFERENCE: 00-28
; CURRENT APPLICATION NUMBER: US/09/540,224
; CURRENT FILING DATE: 2000-03-31
; EARLIER APPLICATION NUMBER: US 60/180,169
; EARLIER FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-540-224-5
```

```
Query Match 18.7%; Score 318.5; DB 4; Length 345;
Best Local Similarity 53.3%; Pred. No. 2.2e-25;
Matches 57; Conservative 14; Mismatches 33; Indels 3; Gaps 1;

Qy 205 VDLRLNDKARYSCPTPRNSVNIREEIKLANVFPFRCILVORCGNCGCTVNNRSC 264
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 236 VDLNLTEBRALYSCPTPRNSVSIREBLKRTDTITFWGCLLVKRCGNCACCLHNCNEQ 295

Qy 265 CNSGKTVKKYHEVLOFEFGHIKRGAKTMAVDIOLDHHERDCIC 311
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 296 CVPSTKYTKYHEVLOLRP---KTGVKGLHKSITDVALHEHBECDVC 339
```

```
RESULT 8
US-09-457-066-43
; Sequence 43, Application us/09457066
; Patent No. 643673
; GENERAL INFORMATION:
```

```
; APPLICANT: Gao, Zeren
; APPLICANT: Hart, Charles E.
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Gilbertson, Debra G.
; APPLICANT: West, James W.
; TITLE OF INVENTION: GROWTH FACTOR HOMOLOG ZVEGF3
; FILE REFERENCE: 98-60
; CURRENT APPLICATION NUMBER: US/09/457,066
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 43
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-457-066-43
```

```
Query Match 18.5%; Score 314.5; DB 4; Length 345;
Best Local Similarity 43.1%; Pred. No. 5.7e-25;
Matches 66; Conservative 19; Mismatches 47; Indels 21; Gaps 3;

Qy 177 VTVLSDISLHMH--KTDRIVNLHMCVYLLT-----VDLRLNDKARYS 218
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 190 VTAFTLEBLIYLEPDRQVDLDSLYKPTWLLGKAFIYKSKSVNLTLEKEVKLYS 249

Qy 219 CTPRNSVNIREEIKLANVFPFRCILVORCGNCGCTVNNRSCNCGKTVKKYHEV 278
| | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 250 CTPRNSVSIREBLKRTDTITFWGCLLVKRCGNCACCLHNCNECCVPRKYTKYHEV 309

Qy 279 OFEFGHIKRGAKTMAVDIOLDHHERDCIC 311
| | | | | : | | | | | : | | | | | : | | | | |
Db 310 QLRP---KTGVKGLHKSITDVALHEHBECDVC 339
```

```
RESULT 9
US-09-540-224-9
; Sequence 9, Application us/09540224
; Patent No. 6468543
; GENERAL INFORMATION:
; APPLICANT: Gilbertson, Debra G.
; APPLICANT: Hart, Charles E.
; TITLE OF INVENTION: METHODS FOR PROMOTING GROWTH OF BONE,
; TITLE OF INVENTION: LIGAMENT AND CARTILAGE USING ZVEGF4
; FILE REFERENCE: 00-28
; CURRENT APPLICATION NUMBER: US/09/540,224
; CURRENT FILING DATE: 2000-03-31
; EARLIER APPLICATION NUMBER: US 60/180,169
; EARLIER FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide
US-09-540-224-9
```

```
Query Match 6.9%; Score 118; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 FEFGIKRGAKTMAVDIOLD 302
| | | | | : | | | | | : | | | | | : | | | | |
Db 2 FEFGIKRGAKTMAVDIOLD 24
```

```
RESULT 10
US-08-915-795-5
; Sequence 5, Application us/08915795
; Patent No. 6235713
```


CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
TISSUE TYPE: Mouse Lung
US-08-915-795-9

Query Match 6.2%; Score 106.5; DB 4; Length 321;
Best Local Similarity 23.8%; Pred. No. 0.0046;
Matches 43; Conservative 26; Mismatches 71; Indels 41; Gaps 8;
QY 150 SCOTIRTKNOTPHVLTWRMILOOSH-----WTVLS-----DISELMHKTRIVNL 196
DB 33 SRMLERSEQOIRAAASLELLQIAHSEDMKMRCKLKSLASMSRSASHSTFA-- 90
QY 197 LMCMYLLTVLDRLNDARKYSCTPRNYSVNIREEL-KLANYVFPFRCLLVORCGNCGC 255
DB 91 --ATFYDTETLKVIDEMQRTQCSPRETCVEVASSELGKTNTFFKPCVNVFRCGCGC-- 146
QY 256 GYVNMRSCTGNSGKT---VKYHEV---LQPEBGHKKRGRATMALVYQIDHHERCQC 309
DB 147 ---NEBGVMCMNTSYISKQLEISVPLTSVP-----ELVPMKIANHTGCKC 191
QY 310 I 310
DB 192 L 192

RESULT 13
US-08-915-795-8
Sequence 8, Application US/08915795
Patent No. 6235713
GENERAL INFORMATION:
APPLICANT: Marc G. ACHEN
APPLICANT: Andrew F. WILKS
APPLICANT: Steven A. STACKER
APPLICANT: Kari ALITALO
TITLE OF INVENTION: GROWTH FACTOR
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eversen, McKewen, Edwards & Lenahan P.L.L.C.
STREET: 1200 G Street, NW, Suite 700
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,795
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 1064/42983
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
TELEX: N/A
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 358 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
TISSUE TYPE: Mouse Lung
US-08-915-795-8

Query Match 6.2%; Score 106.5; DB 4; Length 358;
Best Local Similarity 23.8%; Pred. No. 0.0054;
Matches 43; Conservative 26; Mismatches 71; Indels 41; Gaps 8;
QY 150 SCOTIRTKNOTPHVLTWRMILOOSH-----WTVLS-----DISELMHKTRIVNL 196
DB 38 SRMLERSEQOIRAAASLELLQIAHSEDMKMRCKLKSLASMSRSASHSTFA-- 95
QY 197 LMCMYLLTVLDRLNDARKYSCTPRNYSVNIREEL-KLANYVFPFRCLLVORCGNCGC 255
DB 96 --ATFYDTETLKVIDEMQRTQCSPRETCVEVASSELGKTNTFFKPCVNVFRCGCGC-- 151
QY 256 GYVNMRSCTGNSGKT---VKYHEV---LQPEBGHKKRGRATMALVYQIDHHERCQC 309
DB 152 ---NEBGVMCMNTSYISKQLEISVPLTSVP-----ELVPMKIANHTGCKC 196
QY 310 I 310
DB 197 L 197

RESULT 14
US-09-469-186-1
Sequence 1, Application US/09469186
Patent No. 6361486
GENERAL INFORMATION:
APPLICANT: ACHEN, Marc G.
APPLICANT: STACKER, Steve A.
TITLE OF INVENTION: ANTIBODIES TO TRUNCATED VEGF-D AND USES THEREOF
FILE REFERENCE: ACHEN et al 1064-44660
CURRENT APPLICATION NUMBER: US/09/469,186
CURRENT FILING DATE: 1999-12-21
EARLIER APPLICATION NUMBER: 60/113,254
EARLIER FILING DATE: 1998-12-21
EARLIER APPLICATION NUMBER: 60/134,556
EARLIER FILING DATE: 1999-05-17
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 109
TYPE: PRT
ORGANISM: Homo sapiens
US-09-469-186-1

Query Match 6.0%; Score 102; DB 4; Length 109;
Best Local Similarity 28.7%; Pred. No. 0.0029;
Matches 31; Conservative 16; Mismatches 37; Indels 24; Gaps 5;

OY 210 LINDAKRYSCPRNYSVNIREF--KLANYVFPCLLVQRCGCGCTVMNRCTCNSG 268
 DB 10 IDEEMORQCPRETCEVASSELGKSTTFKPCVNVFRCGGCC---NEESLICMNT 64
 OY 269 KT---VKYHEV---LOPEPGHKKRGRKTMALVDIOLDHERDCDI 310
 DB 65 STYISKQLEFISVPLTSVP-----ELVPVNVNHTGCKL 100

RESULT 15

US-08-999-811-4
 ; Sequence 4, Application US/08999811
 ; Patent No. 5932540
 ; GENERAL INFORMATION:
 ; APPLICANT: HU, JING-SHAN
 ; APPLICANT: ROSEN, CRAIG A.
 ; APPLICANT: CAO, LIANG
 ; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR 2
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX
 ; STREET: 1100 NEW YORK AVENUE
 ; CITY: WASHINGTON
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/999,811
 ; FILING DATE: HEREMITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/207,550
 ; FILING DATE: 8-MAR-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/465,968
 ; FILING DATE: 06-JUN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MARKOWICZ, KAREN R.
 ; REGISTRATION NUMBER: 36,351
 ; REFERENCE/DOCKET NUMBER: 1488.1000004
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)371-2600
 ; TELEFAX: (202)371-2540
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 350 amino acids
 ; TYPE: amino acid
 ; * TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-999-811-4

Query Match 5.7%; Score 96.5; DB 2; Length 350;

Best Local Similarity 19.8%; Pred. No. 0.06; Mismatches 64; Indels 77; Gaps 9;

DB 184 SELAKTKDRIYVNLKMWLLFVDDLRLNDADKRYSCPRNYSVNIREFKLAVVFP-PR 242
 DB 28 ANLMSRTEETIKFAAHYNTFI-LKSIDNEMKTKQCMRREVCIDVGKEFGVANTTFKPP 86
 OY 243 CLVQRCGCGNG-----CGTVMNRCTCNSGK 269
 DB 87 CVSVYRCGCGCNSBGLQCMNTSTSYLSKTLFEITVPLSQPKPTVISFANHSCRCMSKL 146
 OY 270 TV-KYHEVLO-----FEPGHIK-----RGAKTMA 295
 DB 147 DVYRQVSHIIRSLPATLPQCAANKTCPTNYMMNHHICRLAQEDFMFSSDAGDSDTG 206

OY 296 LVDI-----OLDHERDCICSS--RP 315
 DB 207 FHDICGNKELD--ETTCQVCYRAGLR 232

Search completed: June 5, 2003, 12:45:08
 Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 5, 2003, 12:43:54 ; Search time 46 Seconds
(without alignments)
711.460 Million cell updates/sec

Title: US-10-083-853B-1

Perfect score: 1704

Sequence: 1 MGDFPMKTPKAMATYAKID.....DIQLDHERDCICSSRPPR 317

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubppa/PTCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1704	100.0	317	US-10-083-853-2	Sequence 2, Appl1
2	680.5	39.9	956	US-10-000-256A-153	Sequence 153, App
3	672	39.5	917	US-10-001-835-189	Sequence 189, App
4	672.5	39.5	1275	US-10-025-201-3	Sequence 3, Appl1
5	637	37.4	322	US-10-086-623-6	Sequence 6, Appl1
6	637	37.4	322	US-10-260-539-6	Sequence 6, Appl1
7	637	37.4	364	US-10-028-072-186	Sequence 186, App
8	637	37.4	364	US-10-131-049-186	Sequence 186, App
9	637	37.4	364	US-10-123-904-186	Sequence 186, App
10	637	37.4	364	US-10-140-470-186	Sequence 186, App
11	637	37.4	364	US-10-175-746-186	Sequence 186, App
12	637	37.4	364	US-10-176-918-186	Sequence 186, App
13	637	37.4	364	US-10-176-921-186	Sequence 186, App
14	637	37.4	364	US-10-137-865-186	Sequence 186, App
15	637	37.4	364	US-10-142-474-186	Sequence 186, App
16	637	37.4	364	US-10-142-431-186	Sequence 186, App
17	637	37.4	364	US-10-143-114-186	Sequence 186, App
18	637	37.4	364	US-10-140-002-186	Sequence 186, App
19	637	37.4	364	US-10-142-419-186	Sequence 186, App

20	637	37.4	364	9	US-10-123-262-186	Sequence 186, App
21	637	37.4	364	9	US-10-142-423-186	Sequence 186, App
22	637	37.4	364	9	US-10-121-050-186	Sequence 186, App
23	637	37.4	364	9	US-10-141-755-186	Sequence 186, App
24	637	37.4	364	9	US-10-143-032-186	Sequence 186, App
25	637	37.4	364	9	US-10-123-108-186	Sequence 186, App
26	637	37.4	364	9	US-10-123-236-186	Sequence 186, App
27	637	37.4	364	9	US-10-123-261-186	Sequence 186, App
28	637	37.4	364	9	US-10-140-921-186	Sequence 186, App
29	637	37.4	364	9	US-10-140-928-186	Sequence 186, App
30	637	37.4	364	9	US-10-121-045-186	Sequence 186, App
31	637	37.4	364	9	US-10-123-292-186	Sequence 186, App
32	637	37.4	364	9	US-10-123-903-186	Sequence 186, App
33	637	37.4	364	9	US-10-124-819-186	Sequence 186, App
34	637	37.4	364	9	US-10-124-822-186	Sequence 186, App
35	637	37.4	364	9	US-10-140-925-186	Sequence 186, App
36	637	37.4	364	9	US-10-160-498-186	Sequence 186, App
37	637	37.4	364	9	US-10-121-041-186	Sequence 186, App
38	637	37.4	364	9	US-10-121-043-186	Sequence 186, App
39	637	37.4	364	9	US-10-121-047-186	Sequence 186, App
40	637	37.4	364	9	US-10-123-215-186	Sequence 186, App
41	637	37.4	364	9	US-10-123-902-186	Sequence 186, App
42	637	37.4	364	9	US-10-123-908-186	Sequence 186, App
43	637	37.4	364	9	US-10-123-909-186	Sequence 186, App
44	637	37.4	364	9	US-10-123-910-186	Sequence 186, App
45	637	37.4	364	9	US-10-124-813-186	Sequence 186, App

ALIGNMENTS

RESULT 1
US-10-083-853-2
; Sequence 2, Application US/10083853
; Patent No. US20020164709A1
; GENERAL INFORMATION:
; APPLICANT: Affymetrix, Inc
; APPLICANT: Shigeta, Ron T
; APPLICANT: Shani-Rose, Michael A
; TITLE OF INVENTION: Nucleic Acid Encoding Growth Factor Protein
; FILE REFERENCE: 3385.1
; CURRENT APPLICATION NUMBER: US/10/083, 853
; CURRENT FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: USSN 60/272,663
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-083-853-2

Query Match 100.0%; Score 1704; DB 9; Length 317;
Best Local Similarity 100.0%; Pred. No. 4.2e-136;
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGDFPMKTPKAMATYAKIDKMDLTKSFTAKETTRVNRQLTWETKIFATYSDKGL 60
Db 1 MGDFPMKTPKAMATYAKIDKMDLTKSFTAKETTRVNRQLTWETKIFATYSDKGL 60
QY 1 ISRIYELKQIYKKTKNPKIKVWKNRHFSEGIYAAKKHKKYSSSLATREMOIKTT 120
Db 1 ISRIYELKQIYKKTKNPKIKVWKNRHFSEGIYAAKKHKKYSSSLATREMOIKTT 120
QY 121 MRYHLTPVRMAIIKSGNNRMDDEAGNHSQQTITRTKQTPHVLTRWILQSHWTVL 180
Db 121 MRYHLTPVRMAIIKSGNNRMDDEAGNHSQQTITRTKQTPHVLTRWILQSHWTVL 180
QY 181 SDISELHMTDRIVNLIMCNYLLTVLDRLNDAKARYSCTPRYSVNIIBELKLVAVVF 240
Db 181 SDISELHMTDRIVNLIMCNYLLTVLDRLNDAKARYSCTPRYSVNIIBELKLVAVVF 240

Oy 241 PRLIVORCGNCGGTVMRSCTGNSGKTVKCYHEVLQFEPGHIKRRGRATMALVDIQ 300
Db 241 PRLIVORCGNCGGTVMRSCTGNSGKTVKCYHEVLQFEPGHIKRRGRATMALVDIQ 300
Oy 301 LDHHERCDICSSRPFR 317
Db 301 LDHHERCDICSSRPFR 317

RESULT 2
US-10-000-256A-153

; Sequence 153, Application US/10000256A
; Publication No. US2003003983A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0259
; CURRENT APPLICATION NUMBER: US/10/000,256A
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 60/244,782
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 153
; LENGTH: 956
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-000-256A-153

Query Match 39.9%; Score 680.5; DB 9; Length 956;
Best Local Similarity 72.9%; Pred. No. 3.4e-49;
Matches 148; Conservative 5; Mismatches 29; Indels 21; Gaps 5;
Oy 1 MGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 60
Db 671 MGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 730
Oy 61 ISRIYNELKQIYKKKTNPPIKKWVKDMNRHFSKEGIYAAKHKMKYSSSLAIREMOIKTT 120
Db 731 ISRIYNELKQIYKKKTNPPIKKWVKDMNRHFSKEGIYAAKHKMKYSSSLAIREMOIKTT 790
Oy 121 MRHYLTPVRMAIITKSGNNRDMDEAGNHSSQOITTRTKNOTPHVLTIRW---IIQOSHVV 177
Db 791 MRHYLTPVRMAIITKSGNNRDMDEAGNHSSQOITTRTKNOTPHVLTIRW---IIQOSHVV 836
Oy 178 TV---LSDISELMKTDRIVNLL 197
Db 837 SVMRFRLRL-ELRIIPDPAIPL 858

RESULT 3

US-10-001-835-189
; Sequence 189, Application US/10001835
; Patent No. US20020160387A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Cafferey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
; FILE REFERENCE: DEX-0277
; CURRENT APPLICATION NUMBER: US/10/001,835
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/249,997
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 189

; LENGTH: 917
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-835-189

Query Match 39.5%; Score 673; DB 9; Length 917;
Best Local Similarity 94.2%; Pred. No. 1.4e-48;
Matches 131; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Oy 1 MGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 60
Db 779 VGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 838
Oy 61 ISRIYNELKQIYKKKTNPPIKKWVKDMNRHFSKEGIYAAKHKMKYSSSLAIREMOIKTT 120
Db 839 ISRIYNELKQIYKKKTNPPIKKWVKDMNRHFSKEGIYAAKHKMKYSSSLAIREMOIKTT 898
Oy 121 MRHYLTPVRMAIITKSGNN 139
Db 899 MRHYLTPVRMAIITKSGNN 917

RESULT 4
US-10-025-201-3

; Sequence 3, Application US/10025201
; Publication No. US20030003468A1
; GENERAL INFORMATION:
; APPLICANT: Crow, Mary K.
; TITLE OF INVENTION: MARKERS FOR DISEASE SUSCEPTIBILITY AND TARGETS FOR THERAPY
; FILE REFERENCE: 5983/2H567
; CURRENT APPLICATION NUMBER: US/10/025,201
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,673
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 3
; LENGTH: 1275
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank Accession No. US20030003468A1 U09116
; DATABASE ENTRY DATE: 1995-02-02
; RELEVANT RESIDUES: (1)...(1275)
US-10-025-201-3

Query Match 39.5%; Score 672.5; DB 9; Length 1275;
Best Local Similarity 71.4%; Pred. No. 2.3e-48;
Matches 147; Conservative 4; Mismatches 28; Indels 27; Gaps 5;

Oy 1 MGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 60
Db 990 VGKDFMSKTPKAMATKAKIDKMDLIKLSFCTAKETTRVNRQLTWEKIFATYSPDKGL 1049
Oy 61 ISRIYNELKQIYKKKTNPPIKKWVKDMNRHFSKEGIYAAKHKMKYSSSLAIREMOIKTT 120
Db 1050 ISRIYNELKQIYKKKTNPPIKKWVKDMNRHFSKEGIYAAKHKMKYSSSLAIREMOIKTT 1109
Oy 121 MRHYLTPVRMAIITKSGNN--RDMDEAGNHSSQOITTRTKNOTPHVLTIRW---IIQOSHVV 174
Db 1110 MRHYLTPVRMAIITKSGNNRCWRCGEIG-----TLHCWMDCKLAQPIVK 1152
Oy 175 HWYTV---LSDISELMKTDRIVNLL 197
Db 1153 LKSWRFRLRL-ELRIIPDPAIPL 1177

RESULT 5
US-10-086-623-6

; Sequence 6, Application US/10086623
; Patent No. US20020164710A1
; GENERAL INFORMATION:
; APPLICANT: ERIKSSON, ULF

```

APPLICANT: AASE, Karin
APPLICANT: LI, Xuri
APPLICANT: PONTEN, Annica
APPLICANT: TUTELA, Marko
APPLICANT: ALITALO, Kari
APPLICANT: OESSTMAN, Arne
APPLICANT: HELDIN, Carl Henrik
TITLE OF INVENTION: PLATELET DERIVED GROWTH FACTOR D, DNA CODING THEREFOR AND USES THEREOF
FILE REFERENCE: 1064/44833C2
CURRENT APPLICATION NUMBER: US/10/086,623
PRIORITY FILING DATE: 2000-03-04
PRIORITY APPLICATION NUMBER: US 60/107,852
PRIORITY FILING DATE: 1998-11-10
PRIORITY APPLICATION NUMBER: US 60/113,997
PRIORITY FILING DATE: 1998-12-28
PRIORITY APPLICATION NUMBER: US 60/150,604
PRIORITY FILING DATE: 1999-08-26
PRIORITY APPLICATION NUMBER: US 60/157,108
PRIORITY FILING DATE: 1999-10-04
PRIORITY APPLICATION NUMBER: US 60/157,756
PRIORITY FILING DATE: 1999-10-05
PRIORITY APPLICATION NUMBER: US 09/438,046
PRIORITY FILING DATE: 1999-11-10
PRIORITY APPLICATION NUMBER: US 09/691,200
PRIORITY FILING DATE: 2000-10-19
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 322
TYPE: prt
ORGANISM: Homo sapiens
US-10-086-623-6

```

```

Query March 37.4%; Score 637; DB 9; Length 322;
Best Local Similarity 65.4%; Pred. No. 4.5e-46;
Matches 125; Conservative 9; Mismatches 11; Indels 46; Gaps 3;

Qy 173 QSHWTVLSDISELMHKT-----DRIVNLLMC-----M 200
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 132 ETNMESTVSSISGSYNPSVTDPTLADLDKKAIEDTVEDLLKYNPESMEDLENM 191
      :::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy 201 YLFT-----VDDRLNDAKRYSCTPRNVSVNI REELKLVNVFFPRCLV 246
      |||:::|:::|:::|:::|:::|:::|:::|:::|
Db 192 YLDTPRYKGRSYHDKRSKVDDRLNDADKRYSCTPRNVSVNI REELKLVNVFFPRCLV 251
      |||:::|:::|:::|:::|:::|:::|:::|:::|

Qy 247 QRCGNCCGCTVNMWRSCTGNSGKTYKTYHEVLQEPFGHKKRGRKTMALVDIOLDHHR 306
      |||:::|:::|:::|:::|:::|:::|:::|:::|
Db 252 QRCGNCCGCTVNMWRSCTGNSGKTYKTYHEVLQEPFGHKKRGRKTMALVDIOLDHHR 311
      |||:::|:::|:::|:::|:::|:::|:::|:::|

Qy 307 CDCTCSSRPR 317
      |||:::|:::|:::|:::|:::|:::|
Db 312 CDCTCSSRPR 322
      |||:::|:::|:::|:::|:::|:::|

RESULT 6
US-10-260-539-6
; Sequence 6, Application US/10260539
; Publication No. US20030073637A1
; GENERAL INFORMATION:-
; APPLICANT: ERIKSSON, ULF
; APPLICANT: AASE, Karin
; APPLICANT: Li, Xuri
; APPLICANT: PONTEN, Annica
; APPLICANT: UTELLA, Marko
; APPLICANT: ALTITALO, Kari
; APPLICANT: OESTMAN, Arne
; APPLICANT: HELDIN, Carl-Henrik
; TITLE OF INVENTION: PLATELET DERIVED GROWTH FACTOR D, DNA CODING THEREFOR AND USES THEREOF
; FILE REFERENCE: 1064/44833C2
; CURRENT APPLICATION NUMBER: US/10/260.539
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US/10/086.623
; PRIOR FILING DATE: 2000-03-04

```

```

? PRIOR APPLICATION NUMBER: US 60/107,852
? PRIOR FILING DATE: 1998-11-10
? PRIOR APPLICATION NUMBER: US 60/113,997
? PRIOR FILING DATE: 1998-12-28
? PRIOR APPLICATION NUMBER: US 60/150,604
? PRIOR FILING DATE: 1999-08-26
? PRIOR APPLICATION NUMBER: US 60/157,108
? PRIOR FILING DATE: 1999-10-04
? PRIOR APPLICATION NUMBER: US 60/157,756
? PRIOR FILING DATE: 1999-10-05
? PRIOR APPLICATION NUMBER: US 09/438,046
? PRIOR FILING DATE: 1999-11-10
? PRIOR APPLICATION NUMBER: US 09/691,200
? PRIOR FILING DATE: 2000-10-19
? NUMBER OF SEQ ID NOS: 42
? SOFTWARE: SeqIdn version 3.1
? SEQ ID NO 6
? LENGTH: 322
? TYPE: prt
? ORGANISM: Homo sapiens
? US-10-260-539-6

```

Query Match	37.4%	Score 637;	DB 9;	Length 322;
Best Local Similarity	-65.4%;	Pred. No. 4.5e-46;		
Matches 125;	Conservative	9;	Mismatches 11;	Indels 46;
				Gaps 3;

[illegible]

```

RESULT 7
US-10-028-072-186
; Sequence 186, Application US/10028072
; Publication No. US20030004311A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresin, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tunas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OF INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028,072
CURRENT FILING DATE: 2001-12-19
PRIORITY APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26

```

1	PRIOR APPLICATION NUMBER: 60/059113
2	PRIOR FILING DATE: 1997-09-17
3	PRIOR APPLICATION NUMBER: 60/059115
4	PRIOR FILING DATE: 1997-09-17
5	PRIOR APPLICATION NUMBER: 60/059117
6	PRIOR FILING DATE: 1997-09-17
7	PRIOR APPLICATION NUMBER: 60/059122
8	PRIOR FILING DATE: 1997-09-17
9	PRIOR APPLICATION NUMBER: 60/059184
10	PRIOR FILING DATE: 1997-09-17
11	PRIOR APPLICATION NUMBER: 60/059263
12	PRIOR FILING DATE: 1997-09-18
13	PRIOR APPLICATION NUMBER: 60/059352
14	PRIOR FILING DATE: 1997-09-19
15	PRIOR APPLICATION NUMBER: 60/059588
16	PRIOR FILING DATE: 1997-09-19
17	PRIOR APPLICATION NUMBER: 60/059836
18	PRIOR FILING DATE: 1997-09-24
19	PRIOR APPLICATION NUMBER: 60/062250
20	PRIOR FILING DATE: 1997-10-17
21	PRIOR APPLICATION NUMBER: 60/062285
22	PRIOR FILING DATE: 1997-10-17
23	PRIOR APPLICATION NUMBER: 60/062287
24	PRIOR FILING DATE: 1997-10-17
25	PRIOR APPLICATION NUMBER: 60/062814
26	PRIOR FILING DATE: 1997-10-24
27	PRIOR APPLICATION NUMBER: 60/062816
28	PRIOR FILING DATE: 1997-10-24
29	PRIOR APPLICATION NUMBER: 60/063045
30	PRIOR FILING DATE: 1997-10-24
31	PRIOR APPLICATION NUMBER: 60/063082
32	PRIOR FILING DATE: 1997-10-31
33	PRIOR APPLICATION NUMBER: 60/063127
34	PRIOR FILING DATE: 1997-10-24
35	PRIOR APPLICATION NUMBER: 60/063327
36	PRIOR FILING DATE: 1997-10-27
37	PRIOR APPLICATION NUMBER: 60/063329
38	PRIOR FILING DATE: 1997-10-27
39	PRIOR APPLICATION NUMBER: 60/063550
40	PRIOR FILING DATE: 1997-10-28
41	PRIOR APPLICATION NUMBER: 60/063561
42	PRIOR FILING DATE: 1997-10-28
43	PRIOR APPLICATION NUMBER: 60/063704
44	PRIOR FILING DATE: 1997-10-29
45	PRIOR APPLICATION NUMBER: 60/063733
46	PRIOR FILING DATE: 1997-10-29
47	PRIOR APPLICATION NUMBER: 60/063735
48	PRIOR FILING DATE: 1997-10-29
49	PRIOR APPLICATION NUMBER: 60/063738
50	PRIOR FILING DATE: 1997-10-29
51	PRIOR APPLICATION NUMBER: 60/063755
52	PRIOR FILING DATE: 1997-10-17
53	PRIOR APPLICATION NUMBER: 60/064248
54	PRIOR FILING DATE: 1997-11-03
55	PRIOR APPLICATION NUMBER: 60/064809
56	PRIOR FILING DATE: 1997-11-21
57	PRIOR APPLICATION NUMBER: 60/064653
58	PRIOR FILING DATE: 1997-11-24
59	PRIOR APPLICATION NUMBER: 60/065111
60	PRIOR FILING DATE: 1997-11-24
61	PRIOR APPLICATION NUMBER: 60/066770
62	PRIOR FILING DATE: 1997-11-24
63	PRIOR APPLICATION NUMBER: 60/069212
64	PRIOR FILING DATE: 1997-12-11
65	PRIOR APPLICATION NUMBER: 60/069278
66	PRIOR FILING DATE: 1997-12-11
67	PRIOR APPLICATION NUMBER: 60/069334

1	PRIOR FILING DATE: 1997-12-11
2	PRIOR APPLICATION NUMBER: 60/069694
3	PRIOR FILING DATE: 1997-12-16
4	PRIOR APPLICATION NUMBER: 60/072320
5	PRIOR FILING DATE: 1998-01-23
6	PRIOR APPLICATION NUMBER: 60/073612
7	PRIOR FILING DATE: 1998-02-04
8	PRIOR APPLICATION NUMBER: 60/074086
9	PRIOR FILING DATE: 1998-02-09
10	PRIOR APPLICATION NUMBER: 60/0740932
11	PRIOR FILING DATE: 1998-02-09
12	PRIOR APPLICATION NUMBER: 60/077791
13	PRIOR FILING DATE: 1998-03-12
14	PRIOR APPLICATION NUMBER: 60/078910
15	PRIOR FILING DATE: 1998-03-20
16	PRIOR APPLICATION NUMBER: 60/079294
17	PRIOR FILING DATE: 1998-03-25
18	PRIOR APPLICATION NUMBER: 60/079663
19	PRIOR FILING DATE: 1998-02-27
20	PRIOR APPLICATION NUMBER: 60/079728
21	PRIOR FILING DATE: 1998-03-27
22	PRIOR APPLICATION NUMBER: 60/080155
23	PRIOR FILING DATE: 1998-03-31
24	PRIOR APPLICATION NUMBER: 60/081203
25	PRIOR FILING DATE: 1998-04-09
26	PRIOR APPLICATION NUMBER: 60/081229
27	PRIOR FILING DATE: 1998-04-09
28	PRIOR APPLICATION NUMBER: 60/081655
29	PRIOR FILING DATE: 1998-04-14
30	PRIOR APPLICATION NUMBER: 60/081817
31	PRIOR FILING DATE: 1998-04-15
32	PRIOR APPLICATION NUMBER: 60/081818
33	PRIOR FILING DATE: 1998-04-15
34	PRIOR APPLICATION NUMBER: 60/082999
35	PRIOR FILING DATE: 1998-04-24
36	PRIOR APPLICATION NUMBER: 60/083322
37	PRIOR FILING DATE: 1998-04-28
38	PRIOR APPLICATION NUMBER: 60/083545
39	PRIOR FILING DATE: 1998-04-29
40	PRIOR APPLICATION NUMBER: 60/084600
41	PRIOR FILING DATE: 1998-05-07
42	PRIOR APPLICATION NUMBER: 60/084627
43	PRIOR FILING DATE: 1998-05-07
44	PRIOR APPLICATION NUMBER: 60/084637
45	PRIOR FILING DATE: 1998-05-07
46	PRIOR APPLICATION NUMBER: 60/085149
47	PRIOR FILING DATE: 1998-05-12
48	PRIOR APPLICATION NUMBER: 60/085323
49	PRIOR FILING DATE: 1998-05-13
50	PRIOR APPLICATION NUMBER: 60/085338
51	PRIOR FILING DATE: 1998-05-13
52	PRIOR APPLICATION NUMBER: 60/085339
53	PRIOR FILING DATE: 1998-05-13
54	PRIOR APPLICATION NUMBER: 60/085579
55	PRIOR FILING DATE: 1998-05-15
56	PRIOR APPLICATION NUMBER: 60/085657
57	PRIOR FILING DATE: 1998-05-15
58	PRIOR APPLICATION NUMBER: 60/085704
59	PRIOR FILING DATE: 1998-05-15
60	PRIOR APPLICATION NUMBER: 60/086414
61	PRIOR FILING DATE: 1998-05-22
62	PRIOR APPLICATION NUMBER: 60/086430
63	PRIOR FILING DATE: 1998-05-22
64	PRIOR APPLICATION NUMBER: 60/087106
65	PRIOR FILING DATE: 1998-05-28
66	PRIOR APPLICATION NUMBER: 60/088026
67	PRIOR FILING DATE: 1998-06-04
68	PRIOR APPLICATION NUMBER: 60/088730
69	PRIOR FILING DATE: 1998-06-10
70	PRIOR APPLICATION NUMBER: 60/088711
71	PRIOR FILING DATE: 1998-06-10
72	PRIOR APPLICATION NUMBER: 60/088810
73	PRIOR FILING DATE: 1998-06-10

```

; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 19/98-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

```

```

Query Match      37.4%; Score 637; DB 9; Length 364;
Best Local Similarity 65.4%; Pred. No. 5.2e-46;
Matches 125; Conservative 9; Mismatches 11; Indels 46; Gaps 3;

```

```

QY 173 QSHWTVLSDISLHMKT-----DRIVNLMC-----M 200
DB 174 ETWVESVTSISGVSNPSVTDPPTLIADLADKIAEFVTVBLLKYFNPESWQEDLENN 233
QY 201 YLLT-----VDLRLNDAKRYSCPTPNYSVNIREEKLANVFFPRCLV 246
DB 234 YLDTPRYRGSHYDRKSKVDLRLNDAKRYSCPTPNYSVNIREEKLANVFFPRCLV 293
QY 247 QRCGNCGCGTVNWRSCCTNSGKTVKYEVLQFEPGHIKRRGRATMALVLDIQLDHER 306
DB 294 QRCGNCGCGTVNWRSCCTNSGKTVKYEVLQFEPGHIKRRGRATMALVLDIQLDHER 353
QY 307 CDCICSSRPPR 317
DB 354 CDCICSSRPPR 364

```

RESULT 8

```

US-10-121-049-186
; Sequence 186, Application US/10121049
; Publication No. US2003002239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumaas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C17

```

```

; CURRENT APPLICATION NUMBER: US/10/121, 049
; CURRENT FILING DATE: 2002-04-12
; PRIOR Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 186
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-121-049-186

```

```

Query Match      37.4%; Score 637; DB 9; Length 364;
Best Local Similarity 65.4%; Pred. No. 5.2e-46;
Matches 125; Conservative 9; Mismatches 11; Indels 46; Gaps 3;

```

```

QY 173 QSHWTVLSDISLHMKT-----DRIVNLMC-----M 200
DB 174 ETWVESVTSISGVSNPSVTDPPTLIADLADKIAEFVTVBLLKYFNPESWQEDLENN 233
QY 201 YLLT-----VDLRLNDAKRYSCPTPNYSVNIREEKLANVFFPRCLV 246
DB 234 YLDTPRYRGSHYDRKSKVDLRLNDAKRYSCPTPNYSVNIREEKLANVFFPRCLV 293
QY 247 QRCGNCGCGTVNWRSCCTNSGKTVKYEVLQFEPGHIKRRGRATMALVLDIQLDHER 306
DB 294 QRCGNCGCGTVNWRSCCTNSGKTVKYEVLQFEPGHIKRRGRATMALVLDIQLDHER 353
QY 307 CDCICSSRPPR 317
DB 354 CDCICSSRPPR 364

```

RESULT 9

```

US-10-123-904-186
; Sequence 186, Application US/10123904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumaas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C54
; CURRENT APPLICATION NUMBER: US/10/123, 904
; CURRENT FILING DATE: 2002-04-16
; PRIOR Application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 186
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-904-186

```

```

Query Match      37.4%; Score 637; DB 9; Length 364;
Best Local Similarity 65.4%; Pred. No. 5.2e-46;
Matches 125; Conservative 9; Mismatches 11; Indels 46; Gaps 3;

```

```

QY 173 QSHWTVLSDISLHMKT-----DRIVNLMC-----M 200
DB 174 ETWVESVTSISGVSNPSVTDPPTLIADLADKIAEFVTVBLLKYFNPESWQEDLENN 233

```

```

QY 201 YLTT-----VDLRLNDADAKRYSCTPRNVSVNIREELKLANVFFPRCLLV 246
D 234 YLDTPRYRGRSYNDRKSKVDLRLNDADAKRYSCTPRNVSVNIREELKLANVFFPRCLLV 293
QY 247 ORCGNGCGCTVWNRSTCNSGKTVKKYHEVLOFEFGHKKRGRAKTMALVDIOLDHHR 306
D 294 ORCGNGCGCTVWNRSTCNSGKTVKKYHEVLOFEFGHKKRGRAKTMALVDIOLDHHR 353
QY 307 CDCICSSRPPR 317
D 354 CDCICSSRPPR 364

```

RESULT 10

```

US-10-140-470-186
; Sequence 186, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Matanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 186
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-470-186

```

```

Query Match 37.4%; Score 637; DB 9; Length 364;
Best Local Similarity 65.4%; Pred. No. 5, 2e-46;
Matches 125; Conservative 9; Mismatches 11; Indels 46; Gaps 3;

```

```

QY 173 QSHWTVLSDISELMHT-----DRIVNLMC-----M 200
D 174 ETNMEVTSISISGVNSPVTDPTLADALDKIAEFVTEDLKYFNPSWQEDLENN 233
QY 201 YLTT-----VDLRLNDADAKRYSCTPRNVSVNIREELKLANVFFPRCLLV 246
D 234 YLDTPRYRGRSYNDRKSKVDLRLNDADAKRYSCTPRNVSVNIREELKLANVFFPRCLLV 293
QY 247 ORCGNGCGCTVWNRSTCNSGKTVKKYHEVLOFEFGHKKRGRAKTMALVDIOLDHHR 306
D 294 ORCGNGCGCTVWNRSTCNSGKTVKKYHEVLOFEFGHKKRGRAKTMALVDIOLDHHR 353
QY 307 CDCICSSRPPR 317
D 354 CDCICSSRPPR 364

```

```

RESULT 11
US-10-175-746-186
; Sequence 186, Application US/10175746

```

```

; Publication No. US20030027270A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Matanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C153
; CURRENT APPLICATION NUMBER: US/10/175,746
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 186
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-746-186

```

```

Query Match 37.4%; Score 637; DB 9; Length 364;
Best Local Similarity 65.4%; Pred. No. 5, 2e-46;
Matches 125; Conservative 9; Mismatches 11; Indels 46; Gaps 3;

```

```

QY 173 QSHWTVLSDISELMHT-----DRIVNLMC-----M 200
D 174 ETNMEVTSISISGVNSPVTDPTLADALDKIAEFVTEDLKYFNPSWQEDLENN 233
QY 201 YLTT-----VDLRLNDADAKRYSCTPRNVSVNIREELKLANVFFPRCLLV 246
D 234 YLDTPRYRGRSYNDRKSKVDLRLNDADAKRYSCTPRNVSVNIREELKLANVFFPRCLLV 293
QY 247 ORCGNGCGCTVWNRSTCNSGKTVKKYHEVLOFEFGHKKRGRAKTMALVDIOLDHHR 306
D 294 ORCGNGCGCTVWNRSTCNSGKTVKKYHEVLOFEFGHKKRGRAKTMALVDIOLDHHR 353
QY 307 CDCICSSRPPR 317
D 354 CDCICSSRPPR 364

```

```

RESULT 12
US-10-176-918-186
; Sequence 186, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Matanabe, Colin K

```

APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C382
CURRENT APPLICATION NUMBER: US/10/176,918
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 186
LENGTH: 364
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-918-186

Query Match 37.4%; Score 637; DB 9; Length 364;
Best Local Similarity 65.4%; Pred. No. 5.2e-46;
Matches 125; Conservative 9; Mismatches 11; Indels 46; Gaps 3;

QY 173 QSHWTVLSDISELMHKT-----DRIVNLMC-----M 200
Db 174 ETWVESVTSISGVSNSSVTPDTLIADALDKIAEFVTEVLDLKYFNPESQEDLENN 233
QY 201 YLLT-----VDLRINDAKRYSCPTPNYSVNIREEKLANVVFPPCLIV 246
Db 234 YLDTPRYGRSYHDKRSKVDLRINDAKRYSCPTPNYSVNIREEKLANVVFPPCLIV 293
QY 247 ORCGNCGGCTVNMRSCTNSGKTVKYEVLQFEPGHIKRRGRATMALVDIQDHHHR 306
Db 294 ORCGNCGGCTVNMRSCTNSGKTVKYEVLQFEPGHIKRRGRATMALVDIQDHHHR 353
QY 307 CDCICSSRPPR 317
Db 354 CDCICSSRPPR 364

RESULT 13

US-10-176-921-186
Sequence 186, Application US/10176921
Publication No. US20030027276A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C288
CURRENT APPLICATION NUMBER: US/10/176,921
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 186
LENGTH: 364
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-921-186

Query Match 37.4%; Score 637; DB 9; Length 364;
Best Local Similarity 65.4%; Pred. No. 5.2e-46;

Matches 125; Conservative 9; Mismatches 11; Indels 46; Gaps 3;

QY 173 QSHWTVLSDISELMHKT-----DRIVNLMC-----M 200
Db 174 ETWVESVTSISGVSNSSVTPDTLIADALDKIAEFVTEVLDLKYFNPESQEDLENN 233
QY 201 YLLT-----VDLRINDAKRYSCPTPNYSVNIREEKLANVVFPPCLIV 246
Db 234 YLDTPRYGRSYHDKRSKVDLRINDAKRYSCPTPNYSVNIREEKLANVVFPPCLIV 293
QY 247 ORCGNCGGCTVNMRSCTNSGKTVKYEVLQFEPGHIKRRGRATMALVDIQDHHHR 306
Db 294 ORCGNCGGCTVNMRSCTNSGKTVKYEVLQFEPGHIKRRGRATMALVDIQDHHHR 353
QY 307 CDCICSSRPPR 317
Db 354 CDCICSSRPPR 364

RESULT 14

US-10-137-865-186
Sequence 186, Application US/10137865
Publication No. US20030032155A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C154
CURRENT APPLICATION NUMBER: US/10/137,865
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 186
LENGTH: 364
TYPE: PRT
ORGANISM: Homo Sapien
US-10-137-865-186

Query Match 37.4%; Score 637; DB 9; Length 364;
Best Local Similarity 65.4%; Pred. No. 5.2e-46;
Matches 125; Conservative 9; Mismatches 11; Indels 46; Gaps 3;

QY 173 QSHWTVLSDISELMHKT-----DRIVNLMC-----M 200
Db 174 ETWVESVTSISGVSNSSVTPDTLIADALDKIAEFVTEVLDLKYFNPESQEDLENN 233
QY 201 YLLT-----VDLRINDAKRYSCPTPNYSVNIREEKLANVVFPPCLIV 246
Db 234 YLDTPRYGRSYHDKRSKVDLRINDAKRYSCPTPNYSVNIREEKLANVVFPPCLIV 293
QY 247 ORCGNCGGCTVNMRSCTNSGKTVKYEVLQFEPGHIKRRGRATMALVDIQDHHHR 306
Db 294 ORCGNCGGCTVNMRSCTNSGKTVKYEVLQFEPGHIKRRGRATMALVDIQDHHHR 353
QY 307 CDCICSSRPPR 317
Db 354 CDCICSSRPPR 364

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model1

Run on: June 5, 2003, 12:32:48 ; Search time 23 Seconds

(without alignments)
571.652 Million cell updates/sec

Title: US-10-083-853b-1

Sequence: 1 MGKDFMSKTPKMATKAKID.....DIQLDHERDCICSSRPPR 317

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892, seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	658.5	38.6	1259	1	LINI_HUMAN
2	443.5	26.0	1260	1	LINI_NYCCO
3	412.5	24.2	1300	1	POL2_MOUSE
4	125.5	7.4	148	1	VEGH_ORF7
5	116.5	6.8	326	1	VEGD_RAT
6	114.5	6.7	354	1	VEGD_HUMAN
7	106.5	6.2	358	1	VEGD_MOUSE
8	101	5.9	1576	1	YLK3_CAEEL
9	96.5	5.7	419	1	VEGC_HUMAN
10	95.5	5.6	135	1	VEGB_RAT
11	95.5	5.6	207	1	VEGB_MOUSE
12	94	5.5	213	1	PDGA_RABIT
13	92.5	5.4	4092	1	DYHC_YEAST
14	92	5.4	133	1	VEGH_ORF2
15	92	5.4	415	1	VEGC_MOUSE
16	91	5.3	803	1	GYRB_BUCAL
17	89.5	5.3	437	1	CSN2_SCHPO
18	89.5	5.3	704	1	DBF4_YEAST
19	89	5.2	204	1	PDGA_RAT
20	89	5.2	211	1	PDGA_MOUSE
21	89	5.2	211	1	ULP1_YEAST
22	88.5	5.2	621	1	GOG4_HUMAN
23	88.5	5.2	2230	1	GOG4_HUMAN
24	88	5.2	190	1	VEGA_HORSE
25	87.5	5.1	298	1	FORB_METJA
26	87	5.1	1202	1	DPOW_ASCIM
27	87	5.1	1226	1	KFA4_XENLA
28	87	5.1	2349	1	TPR_HUMAN
29	87	5.1	4349	1	DYHC_FUSO
30	86.5	5.1	207	1	VEGB_BOVIN
31	86.5	5.1	1312	1	RA50_YEAST
32	86.5	5.1	3391	1	POLG_DEN26
33	86	5.0	1531	1	YQ38_CAEEL

34	85.5	5.0	738	1	YKF4_YEAST
35	85	5.0	740	1	DDX1_HUMAN
36	85	5.0	1338	1	CIN6_YEAST
37	84.5	5.0	338	1	Y941_METJA
38	84.5	5.0	578	1	OAR2_LYMS
39	84.5	5.0	596	1	PRIM_CLOAB
40	84	4.9	394	1	KIM4_HUMAN
41	84	4.9	1327	1	YNCA_CAEEL
42	83.5	4.9	207	1	VEGB_HUMAN
43	83.5	4.9	486	1	YN53_ROTID
44	83.5	4.9	789	1	A12M_YEAST
45	83	4.9	226	1	PDGA_XENLA

ALIGNMENTS

RESULT 1	LINE1_HUMAN	STANDARD;	PRT; 1259 AA.
AC	P08547;		
DT	01-AUG-1988 (Rel. 08, Created)		
DT	01-AUG-1988 (Rel. 08, Last sequence update)		
DT	01-AUG-1988 (Rel. 08, Last annotation update)		
DE	LINE-1 reverse transcriptase homolog.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE.		
RX	MEDLINE=86230917; PubMed=2423883;		
RA	Hattori M., Kuhara S., Takenaka O., Sakaki Y.;		
RT	"A family of repetitive DNA sequences in primates may be derived		
RT	from a sequence encoding a reverse transcriptase-related protein."		
RL	Nature 321:625-628(1986).		
CC	-1- MISCELLANEOUS: THIS SEQUENCE WAS CONSTRUCTED FROM AN ALIGNMENT OF		
CC	PUBLISHED AND UNPUBLISHED SEQUENCES, DETERMINED IN VARIOUS		
CC	LABORATORIES, BELONGING TO THE LINE-1 FAMILY.		
DR	PIR; A25313; GNMHUL.		
DR	InterPro; IPR005135; Exo_endo_phos.		
DR	InterPro; IPR000300; IPRC.		
DR	InterPro; IPR000477; RVTee.		
DR	Pfam; PF000078; rvt; 1.		
DR	Pfam; PF03372; Exo_endo_phos; 1.		
DR	SMART; SM00128; IPRC; 1.		
KW	RNA-directed DNA polymerase.		
SO	SEQUENCE 1259 AA; 147112 MW; F3BC18A3803919F5 CRC64;		
Query Match	38.6%;	Score 658.5;	DB 1; Length 1259;
Best Local Similarity	69.9%;	Pred. No. 3,4e-45;	
Matches 144;	Conservative 6;	Mismatches 29;	Indels 27; Gaps 5;
QY	1 MGKDFMSKTPKMATKAKIDKMDLKLKSFCTAKETTTIVNROLTWEKIFATYSPDKGL 60		
DB	989 MGKDFMSKTPKMATKAKIDKMDLKLKSFCTAKETTTIVNROLTWEKIFATYSPDKGL 1048		
QY	61 ISRIYVELKQIYKKTNDPKKRVKDMNHFSEKGIYAAKKNKYSSSLAIREWIKTT 120		
DB	1049 ISRIYVELKQIYKKTNDPKKRVKDMNHFSEKGIYAAKKNKYSSSLAIREWIKTT 1108		
QY	121 MRVHLTPVMAIIKKSGNN--RDMDAGNHHSGQITTTKNTQTPHYLTHRW---ILQGS 174		
DB	1109 MRVHLTPVMAIIKKSGNNRNRGCGEIG-----TLHICMDCKLVOP 1151		
QY	175 HWVTV--LSDISELMHKTDRIVNLL 197		
DB	1152 LKSWVRFRDL-ELLEIPDPAIPLL 1176		
RESULT 2	LINE1_NYCCO	STANDARD;	PRT; 1260 AA.


```

AC P08548;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-AUG-1988 (Rel. 08, Last annotation update)
DE LINE-1 reverse transcriptase homolog.
OS Nucleobase coucang (Slow Loris).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Loridae; Nycitebus.
OX NCBI_TaxID=9470;
RN [1]
RP SEQUENCE.
RA MEDLINE=86230917; PubMed=2423883;
RA Hatori M., Kuhara S., Takenaka O., Sakaki Y.;
RT "LI family of repetitive DNA sequences in primates may be derived
RT from a sequence encoding a reverse transcriptase-related protein.";
RL Nature 321:625-628(1986).
CC -I- MISCELLANEOUS: THIS SEQUENCE WAS CONSTRUCTED FROM AN ALIGNMENT OF
CC SIX SEQUENCES, DETERMINED BY THESE AUTHORS BUT NOT SHOWN.
CC BELONGING TO THE LINE-1 FAMILY.
CC PIR; B25313; GNRL1.
DR HSSP; P27695; 1HD7.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR00477; RVTse.
DR Pfam; PF00078; rvc; 1.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR RNA-directed DNA polymerase.
SQ SEQUENCE 1260 AA; 147042 MW; 7A6803DF471F7253 CRC64;

Query Match 26.0%; Score 443.5; DB 1; Length 1260;
Best Local Similarity 50.3%; Pred. No. 6.5e-28;
Matches 95; Conservative 25; Mismatches 52; Indels 17; Gaps 2;

QY 1 MGKDFMSKTPKAMTKAKIDKMDLTKLSFCTAKETTRVNRQLTEWEKIFATYSPDKGL 60
DB 989 LGEIFMRRTPAIEVSKIKHWDLTKLSFCTAKIYKSKASRQSEMEKITAGTISDKGL 1048
QY 61 ISRIYNELKQIKKTKNPIKWKVDMNHRHFSKEGIYAKGMYSSSLAIREMOIKTT 120
DB 1049 ITRHRELKHKIKKTRDPISGMARDLKRNFSEKEDRHITVYKMKSSSLIREMOIKTT 1108
QY 121 MRHILTPRMAIIKKSGNRRMDENGNHNSOOTITRTNQPHVLTFRW---IIQSHWV 177
DB 1109 LRYHLTPRVAHITKSPNQRCWRCGGGKGT-----LLHGMWECPLIRSPWK 1154
QY 178 TVLSDISEL 186
DB 1155 DVMRLRDL 1163

RESULT 3
POL2_MOUSE STANDARD; PRT; 1300 AA.
ID POL2_MOUSE
AC P11369;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Retrovirus-related POL polypeptide [contains: Reverse transcriptase
DE (EC 2.7.7.49); Endonuclease].
GN POL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87064284; PubMed=3023821;
RA Leob D.D., Padgett R.W., Hardies S.C., Shehee W.R., Comer M.B.,
RA Edgell M.H., Hutchison C.A. III;
RT "The sequence of a large LMD element reveals a tandemly repeated 5'
RT end and several features found in retrotransposons.";
RL Mol. Cell. Biol. 6:168-182(1986).
CC -I- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13002; AAA66024.1; ALT_INIT.
DR PIR; B24906; GNMSLL.
DR InterPro; IPR005135; Exo_endo_phos.
DR InterPro; IPR00477; RVTse.
DR Pfam; PF00078; rvc; 1.
DR Pfam; PF03372; Exo_endo_phos; 1.
DR Hydroxylase; Transferase; RNA-directed DNA polymerase; Nuclease;
KW Endonuclease; Polypeptide.
SQ SEQUENCE 1300 AA; 151829 MW; 9EE8108493B89635 CRC64;

Query Match 24.2%; Score 412.5; DB 1; Length 1300;
Best Local Similarity 47.0%; Pred. No. 2.1e-25;
Matches 85; Conservative 30; Mismatches 49; Indels 17; Gaps 2;

QY 2 GKDFMSKTPKAMTKAKIDKMDLTKLSFCTAKETTRVNRQLTEWEKIFATYSPDKGL 61
DB 1017 GEKFIINTAMACAVRSRSHIDKMDLTKLSFCKAKDTVTKRPPDWERIFTEPKSDGLI 1076
QY 62 SRIYNELKQIKKTKNPIKWKVDMNHRHFSKEGIYAKGMYSSSLAIREMOIKTTM 121
DB 1077 SNIKELKQVDFRKNPIKWKSELSKEFSPEYRMAEHLKCCSSSLIREMOIKTTL 1136
QY 122 RYHLTPRMAIIKKSGNRRMDENGNHNSOOTITRTNQPHVLTFRW---IIQSHWV 178
DB 1137 RYHLTPRMAKIKNSGSRRCMCGGKGT-----LLHGMWECRLVQPLMK 1182
QY 179 V 179
DB 1183 V 1183

RESULT 4
VEGH_ORFNT STANDARD; PRT; 148 AA.
ID VEGH_ORFNT
AC P52585;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Vascular endothelial growth factor precursor.
GN A2R.
OS Orf virus (strain NZ7) (OV NZ-7).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Parapoxvirus.
OX NCBI_TaxID=73495;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94076465; PubMed=8254780;
RA Lytle D.J., Fraser K.M., Fleming S.B., Mercer A.A., Robinson A.J.;
RA "Homologs of vascular endothelial growth factor are encoded by the
RT poxvirus orf virus.";
RL J. Virol. 68:84-92(1994).
CC -I- FUNCTION: INDUCES ENDOTHELIAL PROLIFERATION.
CC -I- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S67522; AAB29223.1; -.

```

```

DR HSPB, P15692; 2VPF.
DR InterPro: IPR000072; PD_growth_factor.
DR Pfam: PF00341; PDGF; 1.
DR ProDom: PD001629; PD_growth_factor; 1.
DR SMART: SM00141; PDGF; 1.
DR PROSITE: PS00249; PDGF_1; FALSE_NEG.
DR PROSITE: PS0278; PDGF_2; 1.
DR Mitogen: Growth factor; Glycoprotein; Signal.
DR SIGNAL 1 25
FT CHAIN 26 148
FT DISULFID 46 88
FT DISULFID 77 130
FT DISULFID 81 132
FT DISULFID 71 71
FT DISULFID 80 80
FT CARBOHYD 95 95
SQ SEQUENCE 148 AA; 16078 MW; F0E13BA104CC73F8 CRC64;

Query Match 7.4%; Score 125.5; DB 1; Length 148;
Best Local Similarity 25.5%; Pred. No. 0.0019;
Matches 41; Conservative 18; Mismatches 43; Indels 59; Gaps 6;

QY 193 IVNLICMYLLTVLDRLNDLND-----AKRYSCPTENYSVNIREEU-KLANVVFPP 241
DB 10 VVALLICMYNLPCVCSQSDSPSTNDWMRTLPSKSGCKPDTVVYLGESEYSTNLQVNP 69
QY 242 RCLLVORCGNCG-----CCTVNRSGT-----CNSGKTVKRYHEVLQPEFG 283
DB 70 RCVTVNRSGCGCGGQICITAVETRNVTVSVTVGSSSGTSGVSTN----- 118
QY 284 HIKRGRKATMALVDIQLDHERCDI-----CGRPR 317
DB 119 -----LQISVTEHTKDCIGRTTPTTTRER 147

RESULT 5
VEGD RAT STANDARD; PRT; 326 AA.
AC 035251;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (FIGF).
GN FIGF OR VEGFD.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Yamada Y., Hirata Y., Nezu J., Shimane M.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis and endothelial cell growth, stimulating their proliferation and migration and also has effects on the permeability of blood vessels. May function in the formation of the venous and lymphatic vascular systems during embryogenesis, and also in the maintenance of differentiated lymphatic endothelium in adults. Binds and activates VEGFR-3 (Flt4) receptor (By similarity).
CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- PTM: Undergoes a complex proteolytic maturation which generates a variety of processed secreted forms with increased activity toward VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer linked by disulfide bonds before secretion. The fully processed VEGF-D is composed mostly of two VEGF homology domains (VHDS) bound by non-covalent interactions (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF014827; AAB66557.1; -.
DR HSPB, P15692; 1VPP.
DR InterPro: IPR004153; CXKCX_repeat.
DR InterPro: IPR000072; PD_growth_factor.
DR Pfam: PF00341; PDGF; 1.
DR Pfam: PF03128; CXKCX; 1.
DR ProDom: PD001629; PD_growth_factor; 1.
DR SMART: SM00141; PDGF; 1.
DR PROSITE: PS00249; PDGF_1; 1.
DR PROSITE: PS0278; PDGF_2; 1.
DR Mitogen: Growth factor; Glycoprotein; Signal; Repeat;
KW Cleavage on pair of basic residues; Multigene family.
FT SIGNAL 1 21
FT PROPEP 22 93
FT CHAIN 94 210
FT PROPEP 211 326
FT DOMAIN 227 317
FT REPEAT 227 242
FT REPEAT 263 278
FT REPEAT 282 298
FT REPEAT 306 317
FT DISULFID 116 158
FT DISULFID 147 194
FT DISULFID 151 196
FT DISULFID 141 141
FT DISULFID 150 150
FT CARBOHYD 160 160
FT CARBOHYD 190 190
FT CARBOHYD 292 292
SQ SEQUENCE 326 AA; 37112 MW; 1261AFA373596C00 CRC64;

Query Match 6.8%; Score 116.5; DB 1; Length 326;
Best Local Similarity 24.9%; Pred. No. 0.025;
Matches 45; Conservative 25; Mismatches 70; Indels 41; Gaps 8;

QY 150 SGOITRTKNGPHVLTWMLIQOSH-----WVTVLS-----DISELMKHTDIRVNL 196
DB 38 SRSVLRSSQQRRASTLEELQVASEDWKMKRCKLKSILANDSRSTSRSTFA-- 95
QY 197 LNCMYLLTVLDRLNDLNDKRYSCPTENYSVNIREEU-KLANVVFPPCLLVORCGNCGC 255
DB 96 --ATFYDTETLKVIDEMQRTCCSPRETCEVASSELGTTNFFKPCVNVFRCGCGC-- 151
QY 256 GIVNWRSGTNSGKT--VKYIHEV---IQEPGHIKRGKRAKTMALVDIQLDHERCDI 309
DB 152 ---NEBSVNCMTSTSYISKLFELISVPLTSPV-----ELVVKIANHNGCKC 196
QY 310 I 310
DB 197 L 197

RESULT 6
VEGD HUMAN STANDARD; PRT; 354 AA.
AC 043915;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced growth factor) (FIGF).
GN FIGF OR VEGFD.
OS Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

NCBI_TaxID=9606;
 (1)
 SEQUENCE FROM N.A.
 TISSUE=Lung;
 MEDLINE=9734918; PubMed=9205122;
 Yamada Y., Nezu J.-I., Shimane M., Hirata Y.;
 "Molecular cloning of a novel vascular endothelial growth factor,
 VEGF-D";
 Genomics 42:483-488(1997).
 (2)
 SEQUENCE FROM N.A.
 TISSUE=Lung;
 MEDLINE=98140120; PubMed=9479493;
 Rocchigiani M., Leestring M., Luddi A., Orlandini M., Franco B.,
 Rossi E., Balabio A., Zuffardi O., Oliviero S.;
 "Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1
 between the FIGA and the GRPR genes";
 Genomics 47:207-216(1998).
 (3)
 SEQUENCE FROM N.A.
 MEDLINE=9818549; PubMed=9435229;
 Achen M.G., Jellisch W., Kuuk E., Maekinen T., Vitali A., Wilks A.F.,
 Altalo K., Stacker S.A.;
 "Vascular endothelial growth factor D (VEGF-D) is a ligand for the
 tyrosine kinases VEGF receptor 2 (Flk1) and VEGF receptor 3 (Flt4).";
 Proc. Natl. Acad. Sci. U.S.A. 95:548-553(1998).
 (4)
 PROCESSING, AND SEQUENCE OF 89-94, 100-105 AND 206-213.
 MEDLINE=20011413; PubMed=10542248;
 Stecker S.A., Stenvers K.L., Caesar C., Vitali A., Domagala T.,
 Nice E.C., Roufall S., Simpson R.J., Moritz R., Karpanen T.,
 Altalo K., Achen M.G.;
 "Biosynthesis of vascular endothelial growth factor-D involves
 proteolytic processing which generates non-covalent homodimers";
 J. Biol. Chem. 274:32127-32136(1999).
 (5)
 FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis
 and endothelial cell growth, stimulating their proliferation and
 migration and also has effects on the permeability of blood
 vessels. May function in the formation of the venous and lymphatic
 vascular systems during embryogenesis, and also in the maintenance
 of differentiated lymphatic endothelium in adults. Binds and
 activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.
 (6)
 SUBUNIT: Homodimer; non-covalent and antiparallel.
 (7)
 SUBCELLULAR LOCATION: Secreted.
 (8)
 TISSUE SPECIFICITY: Highly expressed in lung, heart, small
 intestine and fetal lung, and at lower levels in skeletal muscle,
 colon, and pancreas.
 (9)
 -1- PM: Undergoes a complex proteolytic maturation which generates a
 variety of processed secreted forms with increased activity toward
 VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer
 linked by disulfide bonds before secretion. The fully processed
 VEGF-D is composed mostly of two VEGF homology domains (VHDS)
 bound by non-covalent interactions.
 (10)
 -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
 (11)
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

HSP; P15692; IVP.
 DR Genew: HGNC:3708; FIGF.
 DR MIM: 300091; --
 DR InterPro: IPR004153; CXCXC_repeat.
 DR InterPro: IPR000072; PD_growth_factor.
 DR Pfam: PF00341; PDGF_1.
 DR Pfam: PF03128; CXCXC_3.
 DR ProDom: PD001629; PD_growth_factor; 1.
 DR SMART: SM00141; PDGF_1.
 DR PROSITE: PS00249; PDGF_1; 1.
 DR PROSITE: PS50278; PDGF_2; 1.
 KM MitoGen; Growth factor; Glycoprotein; Signal; Repeat;
 Cleavage on pair of basic residues; Multigene family.
 FT SIGNAL 1 21
 FT PROPEP 22 88
 FT CHAIN 89 205
 FT PROPEP 206 354
 FT DOMAIN 222 318
 FT REPEAT 222 237
 FT REPEAT 258 273
 FT REPEAT 277 293
 FT REPEAT 301 318
 FT DISULFID 111 153
 FT DISULFID 142 189
 FT DISULFID 146 191
 FT DISULFID 136 136
 FT DISULFID 145 145
 FT CARBOHYD 155 155
 FT CARBOHYD 185 185
 FT CARBOHYD 287 287
 SO SEQUENCE 354 AA; 40444 MW; 2048D769D75173E CRC64;
 Query Match 6.7%; Score 114.5; DB 1; Length 354;
 Best Local Similarity 24.5%; Pred. No. 0.04; Indels 47; Gaps 9;
 Matches 47; Conservative 27; Mismatches 71;
 QY 145 AGNNH-----SQTTRTNGOTPHVLTNRILQOSH-----WTVL-----SDISE 185
 DB 22 SSNENGVKSSOSTLERSQOIRAASSLLELRITHESEKIMRCRLKSTSDMSRS 81
 QY 186 LMKETDIRVLMCMYLTVDLRDLNDKARVSCPTPNYSVNIRESI-KLVNVFPPRL 244
 DB 82 ASHRSRFPATFYDIEFLAK-----IDEMORTQCSRETCVEVASLSEKSTNFFRPV 137
 QY 245 LVORCGNCCGTVNMRSCNSGKT---VKYHEV---LQFPGHKRGRKATMALVD 298
 DB 138 NVPRCCGCC-----NEESLICMNTSTSYISKOLFISVPLTSVP-----ELVP 180
 QY 299 IOLDHERDCDI 310
 DB 181 VKVANHGCCKL 192
 RESULT 7
 VEGD MOUSE
 ID_VEGD MOUSE STANDARD; PRT; 358 AA.
 AC P97946;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vascular endothelial growth factor D precursor (VEGF-D) (c-fos induced
 growth factor) (FIGF).
 DE FIGF OR VEGFD.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Fibroblast;
 RX MEDLINE=97030254; PubMed=8876195;
 RA Orlandini M., Marconcini L., Ferruzzi R., Oliviero S.;

RT "Identification of a c-fos-induced gene that is related to the
RT platelet-derived growth factor/vascular endothelial growth factor
RT family";
RL Proc. Natl. Acad. Sci. U.S.A. 93:11675-11680(1996).
RN (2)
RN SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RC MEDLINE=97349118; PubMed=9205122;
RA Yamada Y., Nezu J.-I., Shimane M., Hirata Y.;
RT "Molecular cloning of a novel vascular endothelial growth factor,
RT VEGF-D";
RL Genomics 42:483-488(1997).
RN (3)
RN DEVELOPMENTAL STAGE;
RN MEDLINE=9828130; PubMed=9622638;
RA Avantaggiato V., Orlandini M., Acampora D., Oliviero S., Simeone A.;
RT "Embryonic expression pattern of the murine fltg gene, a growth factor
RT belonging to platelet-derived growth factor/vascular endothelial
RT growth factor family";
RL Mech. Dev. 73:221-224(1998).
RN (4)
RN RECEPTOR SPECIFICITY;
RX MEDLINE=21276411; PubMed=11279005;
RA Baldwin M.E., Catmel B., Nice E.C., Roufail S., Hall N.E.,
RA Stenvers K.L., Karikainen M.J., Altalo K., Stacker S.A., Achen M.G.;
RT "The specificity of receptor binding by vascular endothelial growth
RT factor-d is different in mouse and man";
RL J. Biol. Chem. 276:19166-19171(2001).
CC -1- FUNCTION: Growth factor active in angiogenesis, lymphangiogenesis
CC and endothelial cell growth, stimulating their proliferation and
CC migration and also has effects on the permeability of blood
CC vessels. May function in the formation of the venous and lymphatic
CC vascular systems during embryogenesis, and also in the maintenance
CC of differentiated lymphatic endothelium in adults. Binds and
CC activates VEGFR-3 (Flt4) receptor.
CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Highly expressed in fetal and adult lung.
CC -1- DEVELOPMENTAL STAGE: Expressed in a dynamic pattern in several
CC body structures and organs of the embryo such as limb buds,
CC acoustic ganglion, teeth, heart, anterior pituitary as well as
CC lung and kidney mesenchyme, liver, derma, and perosteum of the
CC vertebral column.
CC -1- INDUCTION: By the transcription factor c-fos.
CC -1- PTM: Undergoes a complex proteolytic maturation which generates a
CC variety of processed secreted forms with increased activity toward
CC VEGFR-3 and VEGFR-2. VEGF-D first form an antiparallel homodimer
CC linked by disulfide bonds before secretion. The fully processed
CC VEGF-D is composed mostly of two VEGF homology domains (VHDS)
CC bound by non-covalent interactions (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X99572; CAA67892.1; -
CC EMBL; D89628; BAA14002.1; -
CC HSSP; P15692; 1VPP;
CC PMAA-2DPAGE; P97946; -
CC MGD; MGI:108037; Flgf.
CC InterPro; IPR004153; CXKC repeat.
CC InterPro; IPR000072; PD_growth_factor.
CC Pfam; PF00341; PDGF; 1.
CC Pfam; PF03128; CXKC; 2.
CC ProDom; PD001629; PD_growth_factor; 1.
CC SMART; SM00141; PDGF; 1.
CC PROSITE; PS00249; PDGF_1; 1.
CC PROSITE; PS0278; PDGF_2; 1.

KW Mitogen; Growth factor; Glycoprotein; Signal; Repeat;
KW Cleavage on pair of basic residues; Multigene family.
FT SIGNAL; 1 21
FT PROPEP 22 93
FT CHAIN 94 210
FT PROPEP 211 358
FT DOMAIN 227 323
FT REPEAT 227 242
FT REPEAT 263 278
FT REPEAT 282 298
FT REPEAT 306 323
FT DISULFID 116 158
FT DISULFID 147 194
FT DISULFID 151 196
FT DISULFID 141 141
FT DISULFID 150 150
FT CARBOHYD 160 160
FT CARBOHYD 190 190
FT CARBOHYD 292 292
SQ SEQUENCE 358 AA; 40908 MW; 6636B17BF07037C CRC64;
Query Match 6.2%; Score 106.5; DB 1; Length 358;
Best Local Similarity 23.8%; Pred. No. 0.18;
Matches 43; Conservative 26; Mismatches 71; Indels 41; Gaps 8;
QY 150 SQQTTRTKNQTPHYVLTTRMIIQOSH-----WTVVLS-----DISELMKTRIVNL 196
DB 38 SSMERSBQOIRAASSLEELQIAHSEBWKLRKLKSLASNDSSASRSTFA-- 95
QY 197 LMCWVLLTVDLRLNDADKRYSCPTPNYSVNIREEI-KLVANVFPFRCILVORCGNGCG 255
DB 96 --ATPDTITLKVIBEMWORTGCSFRETCVAVASELCKTNTFFFPCCNVVRGCGC-- 151
QY 256 GVNNRSCTGNSGKT---VKXVHEV---LQEPGHIKRGAKTALVDIJDHHERCDC 309
DB 152 ---NEGVWCMTSTSYISKQFEISVPLTSVP-----ELVVKIANHTGCKC 196
QY 310 I 310
DB 197 L 197
RESULT 8
YIK3_CABEL STANDARD; PRT; 1576 AA.
ID YIK3_CABEL
AC P41951; Q95QP7;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative serine/threonine-protein kinase D1044.3 in chromosome III
DE (EC 2.7.1.1).
GN D1044.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Pauley A., Waterston R.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
RN (2)
RN REVISIONS, AND ALTERNATIVE SPLICING.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a (shown here) and b; are
CC produced by alternative splicing.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC STRONG, TO ZC84.1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U00065; AAK68286.1; -;
DR EMBL; U00065; AAL27237.1; -;
DR HSSP; Q63450.1A06.
DR WormPep; D1044.3a; CE27894.
DR WormPep; D1044.3b; CE29743.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR002899; WRI/EB.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01683; EB; 12.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00289; WRI; 12.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
KW Hypothetical protein; transferase; Serine/threonine-protein kinase;
KW ATP-binding; Alternative splicing.
FT DOMAIN 431 703 PROTEIN KINASE.
FT NP BIND 437 445 ATP (BY SIMILARITY).
FT BINDING 461 461 ATP (BY SIMILARITY).
FT ACT SITE 569 569 BY SIMILARITY.
FT VARSPALIC 1 904 MISSING (IN ISOFORM B).
FT VARSPALIC 905 1012 SPEEDOMLQVSDRPVAVVQSHNGYQKQKTKITYEQLS
ACICLLAENPAAGHVPRTAVVILRDVLQVYNNLV
PTITTVVIRQIHVSALAVITVEYE -> MEETCSSEPK
SNITSPHMKLKEVPPIMICLFFLLQIVISVAGCP
GLTFLPFSNENNOPLCTGPOACGSCYSSGSSRGRTICQYA
STVNNYICCYSTNNQ (IN ISOFORM B).
FT SEQUENCE 1576 AA; 174677 MW; 3A11E573E813498 CRC64;

Query March 5.9%; Score 101; DB 1; Length 1576;
Best Local Similarity 19.6%; Pred. No. 2.8;
Matches 43; Conservative 48; Mismatches 76; Indels 52; Gaps 9;

QY 11 KAWATKAKIDKMDIKLSFCTAK-ETTIRVNRQUTMEKIFATYSFDK-----GLIS 62
DB 334 KNTVTDITDISMDLQCSIFRVLRLMTFFERSRL--KVLPTIEIKFPVAVGNVSVL 391
QY 63 RIVNELKQIYKKKKKNPI---KKV--VKDMNRHSSKGS-----IYAKKH 102
DB 392 TDYQIAKMDNLIENIEIKMDVETVERQAVGEVEYLLDQLGAGAFGCYTVYRKK 451
QY 103 MKKTSSS---LAIREMQIKTTWRYHLTPVMAIIKKSGNRRDMDKA-GNHSSQOITRT 157
DB 452 AQSISENPALKLAKE-----IFMNLNDRSDKSGMDISVKIKIQ 494
QY 158 KNGTPHVLTRHMLQOSHMTVLSDISELMAKTRIVNL 196
DB 495 QLRHPNIVRYRIRIENVNRLIYVMDLQGCSLRDLIITM 533

RESULT 9

VEGC HUMAN STANDARD; PRT; 419 AA.
AC P49767;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vascular endothelial growth factor C precursor (VEGF-C) (Vascular
DE endothelial growth factor related protein) (VRP) (Flt4 ligand) (Flt4-
DE L).
GN VEGFC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCB1 TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 103-120.
RX MEDLINE=96178224; Pubmed=8617204;
RA Joukov V., Pajusola K., Kaipainen A., Chllov D., Lahtinen I., Kuk E.,
RA Saksela O., Kalkkinen N., Aittalo K.,
RT "A novel vascular endothelial growth factor, VEGF-C, is a ligand for
RT the Flt4 (VEGFR-3) and KDR (VEGFR-2) receptor tyrosine kinases.",
RL EMO J. 15:290-298 (1996).
RN [2]
RP ERRATUM.
RX MEDLINE=96203094; Pubmed=8612600;
RA Joukov V., Pajusola K., Kaipainen A., Chllov D., Lahtinen I., Kuk E.,
RA Saksela O., Kalkkinen N., Aittalo K.,
RL EMO J. 15:1751-1751 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Glial tumor;
RX MEDLINE=96312526; Pubmed=8700872;
RA Lee J., Gray A., Yuan J., Luoh S.-M., Avraham H., Wood W.I.,
RT "Vascular endothelial growth factor-related protein: a ligand and
RT specific activator of the tyrosine kinase receptor Flt4.",
RL Proc. Natl. Acad. Sci. U.S.A. 93:1988-1992 (1996).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=97388482; Pubmed=9247316;
RA Fltz L.J., Morris J.C., Towler P., Long A., Burgess P., Greco R.,
RA Wang J., Gassaway R., Nickbarg E., Kovacic S., Charletta A.,
RA Giamotti J., Finerty H., Zollner R., Beter D.R., Leak L.V.,
RA Turner K.J., Wood C.R.,
RT "Characterization of murine Flt4 ligand/VEGF-C",
RL Oncogene 15:613-618 (1997).
RN [5]
RP SEQUENCE OF 32-41; 112-121 AND 228-233, AND MUTAGENESIS OF ARG-227.
RX MEDLINE=97377029; Pubmed=9233800;
RA Joukov V., Sora T., Kumar V., Jeltsch M., Claesson-Welsh L., Cao Y.,
RA Saksela O., Kalkkinen N., Aittalo K.,
RT "Proteolytic processing regulates receptor specificity and activity of
RT VEGF-C.",
RL EMO J. 16:3898-3911 (1997).
RN [6]
RP FUNCTION: Growth factor active in angiogenesis, and endothelial
RP cell growth, stimulating their proliferation and migration and
RP also has effects on the permeability of blood vessels. May
RP function in angiogenesis of the venous and lymphatic vascular
RP systems during embryogenesis, and also in the maintenance of
RP differentiated lymphatic endothelium in adults. Binds and
RP activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.
RN [7]
RP SUBUNIT: Homodimer; non-covalent and antiparallel.
RN [8]
RP SUBCELLULAR LOCATION: Secreted.
RN [9]
RP TISSUE SPECIFICITY: Spleen, lymph node, thymus, appendix, bone
RN marrow, heart, placenta, ovary, skeletal muscle, prostate, testis,
RN colon and small intestine and fetal liver, lung and kidney, but
RN not in peripheral blood lymphocyte.
RN [10]
RP PFM: Undergoes a complex proteolytic maturation which generates a
RP variety of processed secreted forms with increased activity toward
RP VEGFR-3, but only the fully processed form could activate VEGFR-2.
RP VEGF-C first form an antiparallel homodimer linked by disulfide
RP bonds. Before secretion, a cleavage occurs between arg-227 and
RP ser-228 producing an heterotrimer. The next extracellular step
RP of the processing removes the N-terminal propeptide. Finally the
RP mature VEGF-C is composed mostly of two VEGF homology domains
RP (VHds) bound by non-covalent interactions.
RN [11]
RP -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
RN [12]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RN or send an email to license@isb-sib.ch).

EMBL; X94216; CA61907.1; -;
EMBL; U43142; AAA85214.1; -;

```

DR EMBL; U58111; AAB02909.1; -
DR HSSP; P15692; 1VPP.
DR Genew; HGNC:12682; VEGFC.
DR MIM; 601528; -
DR InterPro; IPR004153; CXKC repeat.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR000072; PD_growth_factor.
DR Pfam; PF00341; PDGF_1.
DR Pfam; PF03128; CXKC; 5.
DR PRINTS; PR00438; GFCYSKNOT.
DR PRODOM; PD001629; PD_growth_factor; 1.
DR SMART; SM00141; PDGF; 1.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR Mitogen; Growth factor; Glycoprotein; Signal; Repeat;
KW Cleavage on pair of basic residues; Multigene family.
FT SIGNAL 1 31
FT PROPEP 32 111
FT CHAIN 112 227
FT PROPEP 228 419
FT DOMAIN 280 362
FT REPEAT 280 295
FT REPEAT 304 319
FT REPEAT 328 343
FT REPEAT 347 362
FT DISULFID 131 173
FT DISULFID 162 209
FT DISULFID 166 211
FT DISULFID 156 166
FT DISULFID 165 165
FT CARBOHYD 175 175
FT CARBOHYD 205 205
FT CARBOHYD 240 240
FT MUTAGEN 227
SQ SEQUENCE 419 AA; 46883 MW; 9F598719DB3E014F CRC64;

Query Match 5.7%; Score 96.5; DB 1; Length 419;
Best Local Similarity 19.8%; Pred. No. 1.4;
Matches 41; Conservative 25; Mismatches 64; Indels 77; Gaps 9;

184 SELMKTRDRIIVNLMCMYLLTVDLRNDARKYSCPTPNYSVNIHELKLANVVF-DR 242
97 AAINSTETETIKFAAHNYTEI-LKSIDEMKRTQCMPEVCIDVKEGVALNTFFKRP 155
243 CLLVORCGNGC-----CGTVNWSCTGNSGK 269
156 CVSVYRCGGCCNSEGLQCMNTSTSYLTKTLFEITVPLSQPKVVTISFANHSCRCMSIL 215
270 TV-KKHYEVLQ-----FEGHTR-----RGRAKTMA 295
216 DYYRQVHSLIRSLPATLPQCAANKTCPTNYMNMNHICRLAQEDFMESDAGDSTDG 275
296 LVDI-----QLDHERCICGSS--RP 315
276 FHDICGPKNEID-EETCCVCACAGLRP 301

```

```

RN PARTIAL SEQUENCE FROM N.A. (ISOFORM VEGF-B186).
RP STRAIN=Sprague-Dawley; TISSUE=placenta;
RA Mandiata S.J.; Pepper M.S.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN PARTIAL SEQUENCE FROM N.A. (ISOFORM VEGF-B167).
RC TISSUE=Heart;
RA Weil J.; Eschenhagen T.; Miltmann C.; Scholz H.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Growth factor for endothelial cells. VEGF-B167 binds
CC heparin and neuropilin-1 whereas the binding to neuropilin-1 of
CC VEGF-B186 is regulated by proteolysis (By similarity).
CC -1- SUBUNIT: Homodimer; disulfide-linked. Can also form heterodimer
CC with vegf (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted but remains associated to cells or
CC to the extracellular matrix unless released by heparin (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS: At least 2 isoforms: VEGF-B186 (shown here)
CC and VEGF-B167; are produced by alternative splicing.
CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF032925; AAB86884.1; -
DR EMBL; AF022952; AAB95447.1; -
DR HSSP; P15692; 1VPP.
DR InterPro; IPR000072; PD_growth_factor.
DR Pfam; PF00341; PDGF; 2.
DR PRODOM; PD001629; PD_growth_factor; 1.
DR SMART; SM00141; PDGF; 2.
DR PROSITE; PS00249; PDGF_1; 1.
DR PROSITE; PS50278; PDGF_2; 1.
DR Mitogen; Growth factor; Heparin-binding; Alternative splicing;
KW Multigene family.
FT NON TER 1 1
FT DISULFID 47 91
FT DISULFID 51 93
FT DISULFID 41 41
FT DISULFID 50 50
FT VARSPLIC 106 >135
FT CONFLICT 29 29
FT CONFLICT 37 37
FT CONFLICT 98 98
FT NON TER 135 135
SQ SEQUENCE 135 AA; 15001 MW; A915863D8586F82D CRC64;

Query Match 5.6%; Score 95.5; DB 1; Length 135;
Best Local Similarity 28.8%; Pred. No. 0.43;
Matches 30; Conservative 15; Mismatches 38; Indels 21; Gaps 5;

216 RYSCPTPNYSVNIHELKLANV--FFPCLLVORCGNGCGTVNWSCTGNSGKTVK 273
13 RACQPREVYVPLSHEL-MGNVVKOLVPSCTVYVQCG---GCCPPDGLGECVPIGQHVARM 68
274 VHEVLOFEGHTRKRRRAKTMALVDIQLDHERCICGSSRPPR 317
69 QILMTIY-----PSSQLGEMSLSEHSQCCEC---RPKR 97

```


Db 3257 LEASIEVSKRKYSLIRVNEAIKTEMVNOANLDRSISLVSLFPEK-----3303

Qy 256 GTVNRSTCNSGKTVK-----YHEVLOPEPGHKGAKTALV 297

Db 3304 --ERMINTTKOPSKTOSBLIGNCIISYIETVFGHLNERADMLVIL 3349

RESULT 14

VEGF_OREN2 STANDARD; PRT; 133 AA.

AC P52584;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Vascular endothelial growth factor homolog precursor.

GN A2R.

OS Orf virus (strain NZ2) (OV NZ-2).

OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

OC Parapoxvirus.

OX NCBI_TaxID=10259;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94076465; PubMed=8254780;

RA Lytle D.J., Frazer K.M., Fleming S.B., Mercer A.A., Robinson A.J.,

RT "Homologs of vascular endothelial growth factor are encoded by the

RT poxvirus orf virus";

RL J. Virol. 68:84-92(1994).

CC -1- FUNCTION: INDUCES ENDOTHELIAL PROLIFERATION.

CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; S67520; AAB29220.2; -.

DR HSSP; P15692; IVP.

DR InterPro; IPR000072; PD_growth_factor.

DR Pfam; PF00341; PDGF_1.

DR ProDom; PD001629; PD_growth_factor; 1.

DR SMART; SM00249; PDGF_1.

DR PROSITE; PS00249; PDGF_1; 1.

DR PROSITE; PS50278; PDGF_2; 1.

KM Mitogen; Growth factor; Glycoprotein; signal.

FT SIGNAL 1 20 POTENTIAL.

FT CHAIN 21 133 VASCULAR ENDOTHELIAL GROWTH FACTOR

FT HOMOLOG.

FT DISULFID 36 78 BY SIMILARITY.

FT DISULFID 67 112 BY SIMILARITY.

FT DISULFID 71 114 BY SIMILARITY.

FT DISULFID 61 61 INTERCHAIN (BY SIMILARITY).

FT DISULFID 70 70 INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 85 85 N-LINKED (GLCNAC...) (POTENTIAL).

SO SEQUENCE 133 AA; 14715 MW; 917C0F6883030C39 CRC64;

Query Match 5.4%; Score 92; DB 1; Length 133;

Best Local Similarity 25.0%; Pred. No. 0.81;

Matches 34; Conservative 21; Mismatches 57; Indels 24; Gaps 5;

Qy 193 IYVLMCMYTLTVDDR--LNDADKRYSCPTRNISVNIIR-ELKLANVFFPRCLLYORC 249

Db 8 LVAACLHLYLLNAOSNTKGMSEVLKSGECKRPVIVPSETHPELTSGRFNPCTVLMRC 67

Qy 250 GAGCGCGVNMRSCTCNSGKTVKHYEVLQEPGHIKRGAKTALVDIQDHHRCDC 309

Db 68 GAGCC---NDSLECVTEEVNVMEL-----GASGSGSNGMORLSFVHHKDC 114

Qy 310 -----ICSSRPFR 317

Db 115 RPRFTTTPPTTRPR 130

RESULT 15

VEGC_MOUSE STANDARD; PRT; 415 AA.

AC P97953;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Vascular endothelial growth factor C precursor (VEGF-C) (vascular

DE endothelial growth factor related protein) (VRP) (Flt4 ligand) (Flt4-

DE L).

GN VEGFC.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C;

RX MEDLINE=97164697; PubMed=9012504;

RA Kukk E., Lymbousaki A., Taira S., Kaipainen A., Jeltsch M.,

RA Joukov V., Altalo K.,

RT "VEGF-C receptor binding and pattern of expression with VEGFR-3

RT suggests a role in lymphatic vascular development.";

RL Development 122:3829-3837(1996).

RN [2]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 108-126.

RC STRAIN=BALB/C;

RX MEDLINE=97388482; PubMed=9247316;

RA Fltz L.J., Morris J.C., Towler P., Long A., Burgess P., Greco R.,

RA Wang J., Gassaway R., Nickbarg E., Kovacic S., Charette A.,

RA Giametti J., Finmerly H., Zollner R., Beier D.R., Leak L.V.,

RT "Characterization of murine Flt4 ligand/VEGF-C";

RL Oncogene 15:613-618(1997).

CC -1- FUNCTION: Growth factor active in angiogenesis, and endothelial

CC cell growth, stimulating their proliferation and migration and

CC also has effects on the permeability of blood vessels. May

CC function in angiogenesis of the venous and lymphatic vascular

CC systems during embryogenesis, and also in the maintenance of

CC differentiated lymphatic endothelium in adults. Binds and

CC activates VEGFR-2 (Flk1) and VEGFR-3 (Flt4) receptors.

CC -1- SUBUNIT: Homodimer; non-covalent and antiparallel.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: Expression detected in mesenchymal cells of

CC postimplantation embryos, particularly in the regions where the

CC lymphatic vessels undergo sprouting from embryonic veins, such as

CC the perimetaphoric, axillary and jugular regions, and in the

CC developing mesenterium. Expressed in adult heart, brain, spleen,

CC lung, liver, skeletal muscle and kidney.

CC -1- PTM: Undergoes a complex proteolytic maturation which generates a

CC variety of processed secreted forms with increased activity toward

CC VEGFR-3, but only the fully processed form could activate VEGFR-2.

CC VEGF-C first form an antiparallel homodimer linked by disulfide

CC bonds. Before secretion, a cleavage occurs between arg-227 and

CC ser-228 producing an heterotrimer. The next extracellular step

CC of the processing removes the N-terminal propeptide. Finally the

CC mature VEGF-C is composed mostly of two VEGF homology domains

CC (VHds) bound by non-covalent interactions (By similarity).

CC -1- SIMILARITY: BELONGS TO THE PDGF/VEGF FAMILY OF GROWTH FACTORS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U73620; AAC52984.1; -.

Query Match	Best Local Similarity	Score 92;	DB 1;	Length 415;
Matches	27;	Conservative	18;	Mismatches 42;
			Indels	24;
			Gaps	4
Db	207	LDRLDNDKARYSCTPRNYSVANIREBLKLANVVF-PRCLLVORCGANGCGTVMNRSCTC	265	
Oy	115	LKSIDNEWRKTCQCMREVCIDGKEFGAGATNTFFPBCSVYVRCGGCC-----NSEGLQC	169	
Db	266	NSGKTVKKYHEVLQEPGHIRKRGAKMAL-----VDIQLDHERDCDI	310	
Oy	170	MMTST-----GLSKTLFELTIVPLSGGKRPVITSPANNTSCRM	208	

Search completed: June 5, 2003, 12:42:21
Job time : 25 secs

